ON SEGMENTATION IN MAGNETIC RESONANCE IMAGING

A Thesis Submitted
in Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

by

SUSHMITA DATTA

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

23 SEP 2003 / math
पुरुषोत्तम काली... पुर
भारतीय प्रौद्यो... 145048
क्रमांक सं० A...

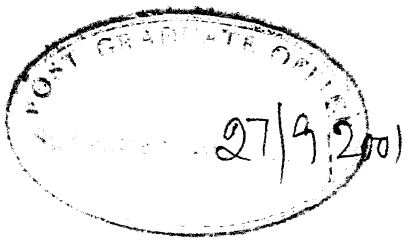
Dedicated

to

the Fond Memory of Dadu

and

Maa & Baba



CERTIFICATE

It is certified that the work contained in the thesis entitled, **"ON SEGMENTATION IN MAGNETIC RESONANCE IMAGING"**, by Sushmita Datta, has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

R.K.S. Rathore
Professor
Department of Mathematics
I.I.T. Kanpur

September, 2001

ACKNOWLEDGEMENTS

I express my sincere gratitude to my thesis supervisor Prof. R.K.S. Rathore for introducing me to the Magnetic Resonance Imaging. I express my thanks to him for his guidance and help during the period of my research. I am also indebted to Auntiji (Dr. (Mrs.) Rathore) for providing me loving, caring and homely treatment during my stay at IIT Kanpur. I always cherish love and affection of Divya and Satya.

I would like to thank Prof. R.K. Gupta, Radiologist, SGPGIMS, Lucknow for providing us with useful and valuable data to carry out this research. I thank him for his guidance and discussions on MRI.

Special thanks are due to Prof. Prabha Sharma, Prof. Debasis Kundu and Prof. Peeyush Chandra, Department of Mathematics, IIT Kanpur, for their valuable suggestions and help.

I will always cherish the company of my friends Anita, Gargi, Mamta, Babu, Sukanta, Sayan, and Shankarshan.

Thanks are due to Mahesh, Kaliprasad, Kalyanaraman, Manish, Muthu, Kailash, Reeta, Sunita, Rajesh, Sanjeev and Monika.

I am thankful to my friends Sharmishthadi, Chaitalidi, Amrapali, Anjali, Ila, Jyoti, Kusum, Leena, Madhuri, Malabika, Moni, Nilam, Radhika, Susmitha, Vijayashree and Trishla for making my stay at IIT Kanpur wonderful and memorable.

I thank Mr. R.K. Jain, Mr. R.K. Srivastava, Mr. N. Islam, Department of Mathematics, IIT Kanpur, for their help during the course of this work.

I am grateful to my friend and colleague S.B. Rao for giving me his kind suggestions, encouragement and help in many ways in the hours of need. I owe to my

friends Mini and Swagata for their love, help and support throughout my stay at IIT Kanpur.

Finally, I wish to express my gratitude to Maa, Baba, Laltoo and Gouri for providing me all the help and encouragement that was required to continue my work. Their faith in me, love and affection made me strong enough to pursue the research.

Sushmita Datta

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF GRAPHS

CHAPTER 0 : INTRODUCTION

Segmentation is a major technique used in the quantitation in magnetic resonance imaging (MRI). This chapter recalls the basics of MRI, and contains a brief survey of the literature on the segmentation of images. This thesis is a study of segmentation of MRI images using different filters, such as Gaussian and Butterworth filters and discrete convolutions with Korovkin, Jackson, generalized Jackson, and De La Vallée-Poussin kernels, followed by maximum likelihood estimation (MLE) based segmentation algorithms developed both for single contrast, as well as multisequence images.

Magnetic resonance imaging is a noninvasive diagnostic technique that produces images of internal body tissues, which is based on nuclear magnetic resonance of atoms within the body induced by the application of radio waves. MRI dominates over other medical imaging modalities due to its excellent soft-tissue contrast, high spatial resolutions and non-invasiveness. MR images are degraded due to, among others, the inhomogeneities in the radio-frequency field (known as bias field intensity) and the presence of multiple tissues within a pixel that leads to a partial volume effect. Presence of many neurological diseases alters the shape and the volume of brain parenchyma and cerebrospinal fluid (CSF) regions. So, their segmentation, and the resulting classification of tissues and pathologies are important in many medical imaging applications. Segmentation helps in the classification and quantitation of normal tissues, as well as different pathologies. Segmentation also helps in obtaining computerized anatomical structures, and in guiding surgery.

An image segmentation is defined as an exhaustive partitioning of an input image into regions, each of which is considered to be homogeneous with respect to some image property of interest (e.g., intensity, color or texture). Segmentation is one of the major steps in the analysis of processed image data. It has a different significance in different areas. Like its main aim in aerial photography is to divide a photographic or geographical or satellite image into parts that have a strong

correlation with objects or areas of the real world contained in the image. In medical images, like ultrasound, computerized tomography, magnetic resonance images, the main aim of segmentation is to separate the region of interest from rest of the image. Segmentation of magnetic resonance images is complex due to various kinds of degradations, and as such no single method seems to be applicable to all kind of images.

In MRI, Segmentation helps in the classification of normal tissues along with the pathologies present within them. It also helps in the quantitative analysis that provides a better knowledge of the growth and diagnosis of pathology during an interval of time and this information could be used in further treatment. It is useful in the understanding of physiology of brain and other parts of the body by enabling a study of the various chemical and biological environments of the tissues. Misclassification of pixels into different tissue classes due to the presence of noise could give improper segmentation which consequently would lead to poor tissue quantitation. The proper choice of filter to reduce noise would help in minimizing these misclassifications and the corresponding errors.

0.1. BASICS OF MAGNETIC RESONANCE IMAGING

Precessional frequency: Magnetic nuclei, when placed in an external magnetic field, $\mathbf{B_0}$, possess a unique precessional motion with Larmor frequency, f_L, which is directly proportional to both the strength of the external magnetic field and gyromagnetic ratio, γ, of the magnetic nuclei and is given by

$$f_L = \frac{\gamma}{2\pi}\mathbf{B}_0 .$$

The gyromagnetic ratio is defined as the ratio of the magnitude of the magnetic moment of the nucleus and the magnitude of its spin angular momentum.

Resultant Magnetization Vector: Resultant magnetization vector **M** of nuclei is the net sum of all the magnetic moment vectors which align in the direction of external magnetic field, $\mathbf{B_0}$, on the application of $\mathbf{B_0}$.

Radio-frequency Magnetic Field: The perpendicular radio-frequency field is produced using a brief application of an alternating electric current through a coil.

Whenever radio-frequency magnetic field $\mathbf{B_1}$, perpendicular to **M,** is imposed, it results in a precession of **M** around $\mathbf{B_1}$ tipping away its orientation from its original position, due to which **M** acquires a component perpendicular to that of $\mathbf{B_0}$ causing it to precess around $\mathbf{B_0}$ also. **M** just wobbles about its original direction as long as $\mathbf{B_1}$ continues to change at an arbitrary frequency. The angle of tip continues to become large in resonance, which is caused when the frequency of $\mathbf{B_1}$ equals the Larmor frequency due to the field $\mathbf{B_0}$. Besides its longitudinal component $\mathbf{M_L}$, the transverse component $\mathbf{M_T}$ of **M** is produced as **M** is tipped from its original orientation (Figure 0.1) following radio-frequency pulse (C.L. Partain, R.R. Price, J.A. Patton, M.V. Kulkarni and A.E. James Jr. [1988]).

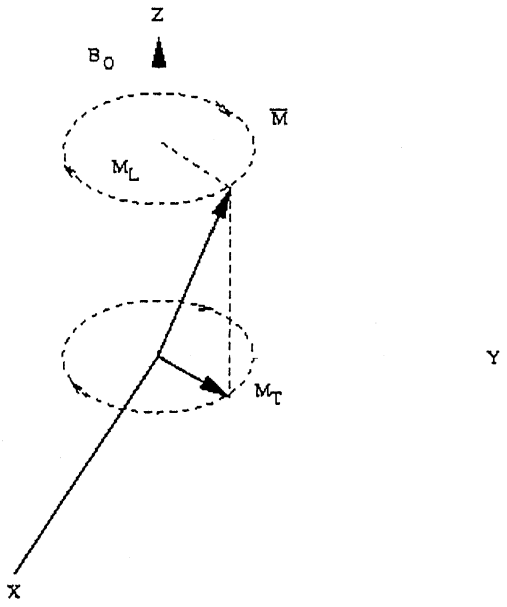


Figure 0.1: The Larmor precession geometry **M** traces out a circle parallel to xy plane and its component $\mathbf{M_T}$ traces out a similar circle in the xy plane, while other component $\mathbf{M_L}$ remains fixed.

A sustained pulling of **M** away from zero can occur only if $\mathbf{B_1}$ is always perpendicular to **M** which makes $\mathbf{B_1}$ to rotate in the same direction as **M** at the rate f_L (Figure 0.2). With $\mathbf{B_1}$ rotating at this rate, **M** remains perpendicular to it while Larmor precessing around it at the rate:

$$f_{sf} = \frac{\gamma}{2\pi}\mathbf{B}_1,$$

where f_{sf} is the spin flip precession rate.

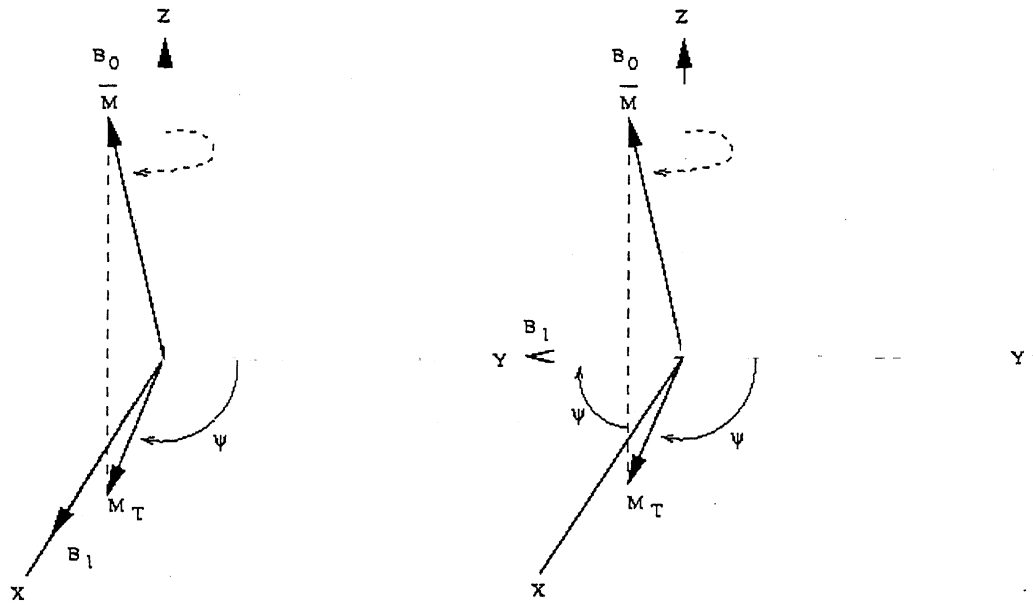


Figure 0.2: *Left*: $\mathbf{B}_1$ is stationary along the x-axis. It pulls **M** clockwise away from z-axis, but $\mathbf{B}_0$ also pushes **M** toward the xz plane. $\mathbf{M}_T$ has rotated through the angle ψ, changing the angle between $\mathbf{B}_1$ and **M**. *Right*: $\mathbf{B}_1$ has been rotated as the same rate as $\mathbf{M}_T$; this keeps the angle between **M** and $\mathbf{B}_1$ fixed, and hence the force on **M** is the same, pulling **M** away from $\mathbf{B}_0$ even further.

During relaxation, $\mathbf{B}_1$ is absent, $\mathbf{M}_L$ grows from its spin-flipped value toward an equilibrium value of **M** in a fully relaxed sample. The rate of relaxation is proportional to the difference between $\mathbf{M}_L$'s current value and its equilibrium value at full relaxation. The time constant of this relaxation is called T_1, and the growth rate of $\mathbf{M}_L$ is referred to as T_1 relaxation, also termed as spin-lattice relaxation. At the same time, $\mathbf{M}_T$ decreases to zero as it rotates at the rate f_L around xy plane. The rate of decrease of $\mathbf{M}_T$ to its equilibrium value of zero is proportional to the current magnitude of $\mathbf{M}_T$. The time constant of this relaxation is called T_2, and the decay rate of $\mathbf{M}_T$ is referred to as T_2 relaxation (also termed as spin-spin relaxation) (C.L. Partain, R.R. Price, J.A. Patton, M.V. Kulkarni and A.E. James Jr. [1988]).

The maximum signal in the receiver coil, following 90° pulse starts decreasing owing to relaxation. This property of signal is known as the free induction decay (FID). Pulse cycle is a repeating unit, which is composed of a series of one or more radio-frequency pulses with a measurement of one or more MR signals. Pulse sequence is a series of pulse cycles. TR (Time to Repeat) is the time interval between two successive pulse cycles, measured in milliseconds. TE (Time to Echo) is the time from one pulse to the measurement of MR signal, which is also measured in milliseconds. T_1 weighted images owe their contrast mainly to T_1 relaxation properties of the tissues. They use a short TE, and a short TR. T_2 weighted images, with contrast predominantly due to the T_2 relaxation properties of the tissues, are produced by using long TE and a long TR. Proton Density weighted image is produced using long TR, with short TE (A.L. Horowitz [1989]). Their contrast depends on the concentration of water molecules in the tissues.

Magnetization transfer (MT) contrast imaging is used to suppress background tissues to enhance vessels and certain disease processes. In this process, prior to the conventional pulse sequence procedure, a pre-saturation pulse is used to saturate the bound protons and promote an exchange of some of this saturated magnetization onto the free protons. This results in reduced signal from the free protons, e.g., gray and white matter lose 30-40% of their signal when an MT pulse sequence is utilized.

The spin echo (SE) sequence can be written as (90° - TE/2 - 180° - TE/2 - measure - TD)$_n$ (TD is a delay time). Since initially **M** is aligned with $\mathbf{B_0}$, $\mathbf{M_L} = \mathbf{M}$ and $\mathbf{M_T} = 0$. Following the 90° pulse, $\mathbf{M_L} = 0$ and $\mathbf{M_T} = \mathbf{M}$, with both T_1 and T_2 relaxation beginning and allowed to continue (as FID) for a time TE/2. The relaxation equations are given as:

$$M_L(t) = M(1 - e^{-t/T_1}) + M_L(0)e^{-t/T_1}$$

$$M_T(t) = M_T(0)e^{-t/T_2} \qquad \text{(True } T_2 \text{ decay)}$$

$$M_T(t) = M_T(0)e^{-t/T_2^*} \qquad \text{(FID decay).}$$

Here $M_L(0)$ is the value of $\mathbf{M_L}$ following a radio-frequency pulse. $\mathbf{M_L}$ is completely relaxed at the start of the first sequence, but only partially relaxed at the end of the sequence, which is a recovery from the 90° and 180° pulses. $\mathbf{M_L}$ ends with same

partially relaxed value on second and subsequent sequences with which it started after the first sequence and the signal intensity can be written as:

$$I_{SE} = kM(1 + e^{-TR/T_1} - 2e^{-(TR-TE/2)/T_1})e^{-TE/T_2}.$$

0.2. SEGMENTATION

In magnetic resonance imaging, the segmentation methods are broadly classified into two categories based on a single image contrast and multi-sequence images (L.P. Clarke, R.P. Velthuizen, S. Phuphanich, J.D. Schellenberg, J.A. Arrington and M. Silbiger [1993], L.P. Clarke, R.P. Velthuizen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher and M.L. Silbiger [1995]). Methods corresponding to a single image are further classified into two categories - edge based, and region based methods. Edge based methods consist of edge detection, boundary tracing and thresholding; whereas region based methods are seed growing, merge and split algorithms and template matching. For multi-sequence images, supervised and unsupervised algorithms are mentioned in the literature.

Edges in camera photographs are due to a change in some physical properties like surface reflectance, the geometry and the intensity of an object (A.K. Jain [1995]). The well- known edge detector, Sobel operator, is defined as the square root of the sum of squares of gradients in x- and y-directions. Edges are also detected by finding zero-crossings in the convolutions of the image with Laplacian-of-Gaussian masks (A. Huertas and G. Medioni [1986]), or by Marr-Hildreth operator (R.M. Haralick [1984], E.C. Hildreth [1983]) which can be approximated by Difference-of-Gaussians that seems to be good to obtain the edges for large regions. Marr-Hildreth operator is used along with morphological filtering (M. Bomans, K-H. Höhne, U. Tiede and M. Riemer [1990]), which requires manual labeling and editing of the regions to generate satisfactory 3-D displays. The edge detection algorithms generally suffer from incorrect detection of edges due to noise, over segmentation due to the presence of inhomogeneity within the region and under segmentation due to inaccurate boundaries between the adjacent regions.

Boundaries are obtained by joining the edge element at a certain pixel to that present in its first order neighborhood or in second order neighborhood. First order neighborhood consists of four horizontal and vertical pixels, while second order neighborhood has first order neighborhood along with four diagonal pixels. In contour following algorithms (G.P. Ashkar and J.W. Modestino [1978]), the pixel is viewed as having a square-shaped boundary, and the boundary is traced by following the edge pixels using graph searching methods, where a pixel acts as a node in a graph (U. Montanari [1971]). In some cases, boundaries are obtained using Fourier descriptors (A. Chakraborty, L.H. Staib and J.S. Duncan [1996], L.H. Staib and J.S. Duncan [1992]) and B-splines (A.K. Jain [1995]).

Thresholding produces a segmentation, which gives all the pixels, either belonging to objects of an interest or background of an image. This is mainly based on the shape of the histogram of image data, such as number of non-overlapping peaks present in histogram plot according to which single or multiple thresholds can be predefined and the thresholding is known as a simple thresholding or multiple thresholding (J.C. Russ [1992]). Optimal thresholding is obtained by the approximation of the histogram of an image using a weighted-sum of two or more probability densities with normal distributions. These algorithms in magnetic resonance imaging are also affected by inhomogeneity present within region of interest.

Seed growing is a region-based method, which selects the seed and predefined threshold, depending on the region of interest. It then examines the pixels surrounding the seed on the basis of intensity or some textural properties of seed and includes the neighboring pixel into the region if it comes within a pre-decided threshold (R.C. Gonzalez and R.E. Woods [1999]). The inclusion of pixels is mainly according to first and second order neighborhood. Each new added pixel then becomes a seed and the process continues till there is no inclusion of new pixel into the region. This method is helpful in segmenting binary images and it is mainly used as a post processing technique to operate on segmented image obtained from some other methods.

Region splitting and merging (R.C. Gonzalez and R.E. Woods [1999]) is an iterative procedure which starts with the splitting of an image into disjoint sets and then the merging takes place and stops when no further splitting and merging is possible. Splitting of a region R on an image into its n subregions $R_1, R_2, \ldots, R_n$, and merging between any two regions are done based on the following conditions:

(1) $\bigcup_{i=1}^{n} R_i = R$,

(2) R_i is a connected region, $i = 1, 2, \ldots, n$,

(3) $R_i \cap R_j = \phi$ for all i and j, $i \neq j$,

(4) $P(R_i)$ = TRUE for $i = 1, 2, \ldots, n$,

(5) $P(R_i \cup R_j)$ = FALSE for $i \neq j$,

Here $P(R_i)$ is a logical predicate over the points in set R_i and ϕ is the null set.

Template matching is based on the matching of a small section, either of certain tissue or abnormality, with the given image (A.K. Jain [1995]). This generally gives location of abnormality and fails to provide any kind of distinction either in terms of boundary or in terms of regions between pathology and normal tissues.

Supervised methods require initial training data to begin with, which is mainly based on rough segmentation or prior knowledge of some features of an image. One of the famous non-parametric supervised classifiers, k-Nearest neighbor (M. Vaidyanathan, L.P. Clarke, C. Heidtman, R.P. Velthuizen and L.O. Hall [1997], L.P Clarke, R.P. Velthuizen, S. Phuphanich, J.D. Schellenberg, J.A. Arrington and M. Silbiger [1993]), uses a sample set of pixel vectors Z and assigns them to different tissue classes. The unlabeled training pixel vectors X are assigned according to the Euclidean distance between X and the labeled pixel vectors Z. The posteriori probability for X is then estimated from the frequency of the labels k-NNs and the tissue class label is assigned to X based on the maximum posteriori probability. Fuzzy c-means (D.L. Pham and J.L. Prince [1999]) is a well-known supervised algorithm for segmentation.

Unsupervised methods automatically find the structure or pattern present within the image and one such well known method is clustering. It is a region based

method which uses the spatial co-ordinates of the centroids of features which are usually gray level intensity, texture, etc., and sorts through them to locate the nearest neighbor features for each feature present. It is based on center-to-center distances and then constructs a distribution image of the intensity of nearest neighbor distances. Some unsupervised techniques could be further processed via multi-spectral analysis along with region analysis (M.C. Clark, L.O. Hall, D.B. Goldgof, R. Velthuizen, F.R. Murtagh and M.S. Silbiger [1998]) to get the required segmentation.

Some algorithms (M.C. Clark, L.O. Hall, D.B. Goldgof, R. Velthuizen, F.R. Murtagh and M.S. Sibiger [1998]) integrate knowledge-based techniques with multi-spectral analysis. Initial segmentation is obtained by an unsupervised clustering algorithm following the multi-spectral histogram analysis and classification of segments based on region analysis. Knowledge of tissue intensity properties and intensity inhomogeneities are used to correct and segment MR images via expectation-maximization algorithm (W.M. Wells III, W.E.L. Grimson, R. Kikinis and F.A Jolesz [1996]).

Segmentations based on normal mixture models of different kinds are also mentioned in the literature. The pixel intensity either follows univariate normal distribution in case of a single image or multivariate normal distributions in case of multiple images.

In Y. Zhang, M. Brady and S. Smith [2001], it is assumed that the pixel intensity follows a Gaussian distribution with parameters $\theta_l = \{\mu_l, \sigma_l\}$, given the class label $x_i = l$,

$$p(y_i \mid x_i) = \frac{1}{\sigma_l \sqrt{2\pi}} \exp\left\{ - \frac{(y_i - \mu_l)^2}{2\sigma_l^2} \right\},$$

and the joint likelihood probability, based on the conditional independence assumption of **y** is given by

$$P(\mathbf{y} \mid \mathbf{x}) = \prod_{i \in S} p(y_i \mid x_i) = \frac{1}{(2\pi)^{N/2}} \exp\left[\sum_{i \in S} \left(- \frac{(y_i - \mu_{x_i})^2}{2\sigma_{x_i}^2} - \log \sigma_{x_i} \right) \right],$$

where S is a two dimensional rectangular lattice and each pixel is characterized by an intensity value y_i and the labeling of S is denoted by $\mathbf{x}$, where $x_i, (i \in S)$ is the corresponding class label of pixel i. The random field $X = \{X_i, i \in S\}$ is an underlying Markov random filed (MRF) assuming values in a finite state space L with probability distribution $P(\mathbf{x}) = Z^{-1} \exp(-U(\mathbf{x}))$ (J. Besag [1974]), where $Z = \sum_{x \in X} \exp(-U(\mathbf{x}))$ is a normalizing constant called the partition, and $U(\mathbf{x})$ is an energy function of the form $U(\mathbf{x}) = \sum_{c \in C} V_c(\mathbf{x})$ which is a sum of clique potentials $V_c(\mathbf{x})$ over all possible cliques C. A clique C is defined as a subset of sites in S in which every pair of distinct sites are neighbors, except for single-site cliques (S. Geman and D. Geman [1984]). Labeling of class in an image is done according to the maximum a posteriori (MAP) criterion,

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x} \in X} \{P(\mathbf{y}/\mathbf{x}) P(\mathbf{x})\}.$$

A. Lundervold and G. Storvik [1995] propose a method for segmentation of brain parenchyma and cerebrospinal fluid, where head and brain are divided into four major regions and seven different tissue types. Each region is associated with a finite mixture of its constituent tissue types and each tissue is modeled by a multivariate Gaussian distribution. In this paper, initial estimates of tissue parameters, means and covariances of mixture are obtained by using k-means clustering algorithm. The boundary of each region is extracted using contour detection algorithm followed by the classification and distribution of tissues in each of the regions. This segmentation method is restricted to images where the brain parenchyma and CSF spaces form connected regions.

The process of segmentation is to find x which represents the correct tissue class at each voxel site given by image y (J.C. Rajapakse, J.N. Giedd and J.L. Rapoport [1997]). To find the maximum a posteriori (MAP) estimation from the image data, that is, if $x = x^*$ represents the optimal segmentation

$$x^* = \arg\max_{x \in \Omega} p(x \mid y),$$

where Ω is the set of all possible segmentations and $p(x|y)$ is the posterior density of the segmentation x given the image y. From Bayes' theorem, $p(x|y) \propto p(x,y) = p(y|x)$,

$p(x)$ due to the independence of prior probability of image $p(y)$, where $p(y|x)$ is the conditional density of a pixel intensity y given segmentation x as:

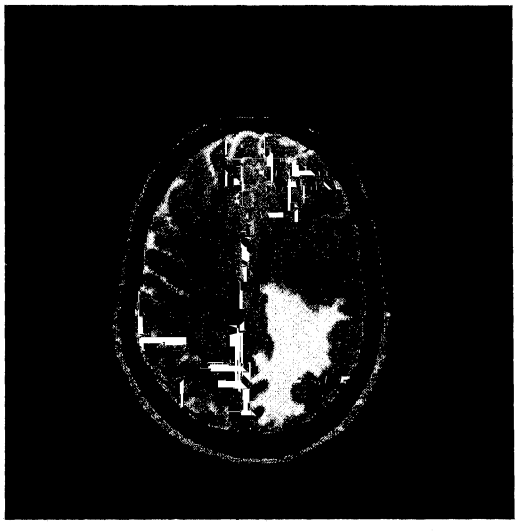
$$p(y \mid x) = \prod_{k} \prod_{i \in R_k} \frac{1}{\sqrt{2\pi}\sigma_{k,i}} \exp\left(-\frac{1}{2}\left(\frac{y_i - \mu_{k,i}}{\sigma_{k,i}}\right)^2\right),$$

where R_k denotes the region or the set of all voxel sites belonging to tissue class k, $\mu_{k,i}$ and $\sigma_{k,i}$ represent the mean intensity and standard intensity of the class k at pixel site i.

Comparisons of different segmentation algorithms according to the performance on intensity inhomogeneities within each tissue region, along with their stability and applications (M. Vaidyanathan, L.P. Clarke, C. Heidtman, R.P. Velthuizen, and L.O. Hall [1997], L.P. Clarke, R.P. Velthuizen, S. Phuphanich, J.D. Schellenberg, J.A. Arrington, and M. Silbiger [1993], D.L. Pham, and J.L. Prince [1999]), like maximum likelihood method, clustering, artificial neural networks (S.C. Amartur, D. Piraino and Y. Takefuji [1992], W.E. Reddick, J.O. Glass, E.N. Cook, T.D. Elkin and R.J. Deaton [1997]) are also reported in the literature.

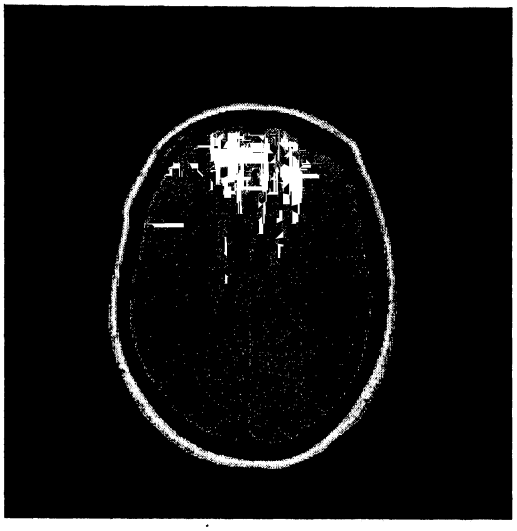
0.3. TEST IMAGE SETS

All numerical simulations in this thesis are done on the data obtained from Department of Radio-diagnosis, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow. Magnetic Resonance Imaging of patients are performed on a 1.5 T super-conducting system using circularly polarized head coil. Conventional Proton density (PD), T_2 (TR/TE/1,2/n = 2200/12, 80/1) and T_1 (1012/14/2) weighted SE images are acquired in axial plane using 256x256 matrix size, 0.1mm inter slice gap and 5mm slice thickness. We have considered three sets of images for our numerical simulations in subsequent chapters. Figure 0.3 shows set I which consists of T_2 weighted image and Figure 0.4 shows set II which consists of Proton density, T_2 and T_1 weighted images. Set III consists of Proton density, T_2, T_1 and MT SE T_1 weighted images shown in Figure 0.5. Magnetization transfer T_1 weighted imaging is performed with identical parameters as for T_1 weighted image except for an off-resonance pulse. Sets I and II are the images of different cross-sections of a patient with edema located in the region of brain parenchyma. Set III is a case of post-traumatic epilepsy with perilesional gliosis.



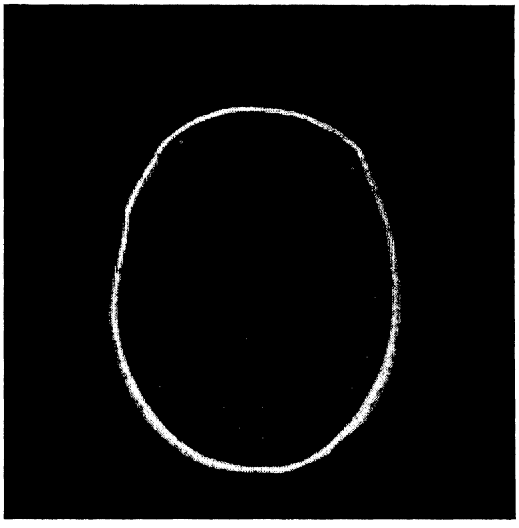
T_2 weighted Image

Figure 0.3: Test Image Set I



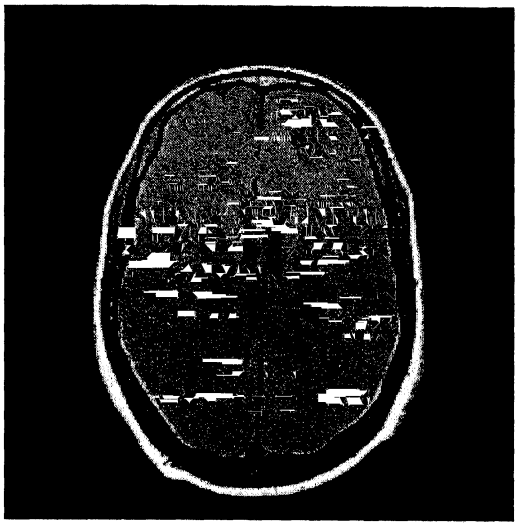
Proton Density weighted Image

T_2 weighted Image

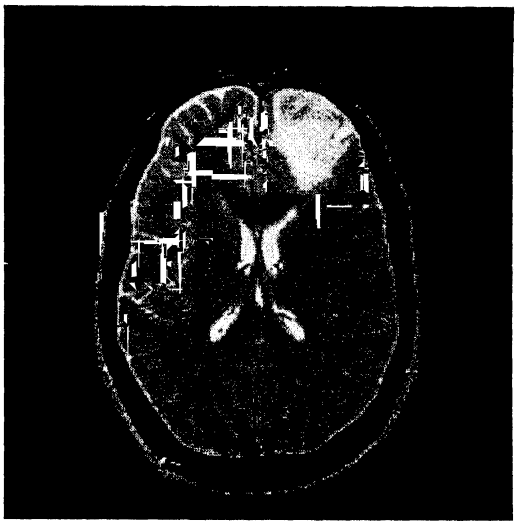


T_1 weighted Image

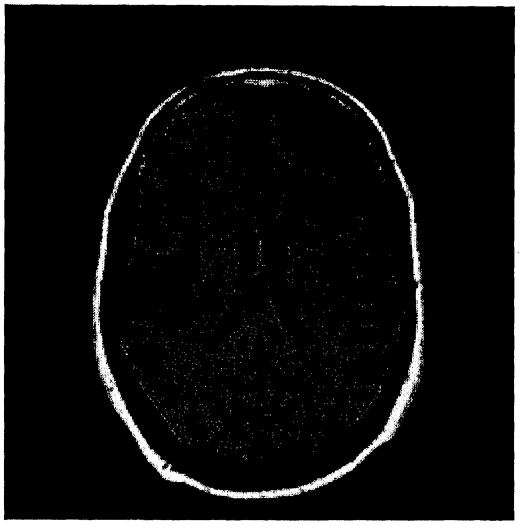
Figure 0.4: Test Image Set II



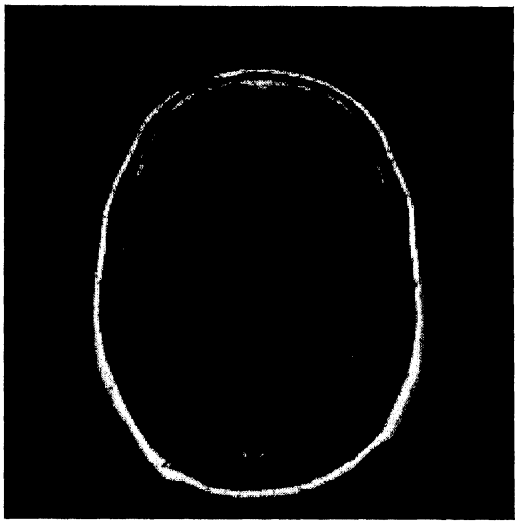
Proton Density weighted Image



T_2 weighted Image



T_1 weighted Image



MT SE T_1 weighted Image

Figure 0.5: Test Image Set III

0.4. CONTENTS OF THE THESIS

A detailed study of the degree of approximation properties of the Marr-Hildreth operator and other related convolutions with the Gaussian function and the resulting segmentations has been done in **Chapter I.**

Asymptotic estimates for Marr-Hildreth operator are deduced in section 1.3 which are as follows:

Theorem 1.3.1.

If $I(X)$ is 4-differentiable at a point X,

$$\lim_{\sigma\to 0}\sigma^{-2}\left\{I(X)\otimes\nabla^2 G(X,\sigma)-\nabla^2 I(X)\right\}=\nabla^4 I(X)/2,$$

and this relation holds uniformly in $X\in S$ if $I(X)$ is uniformly 4-differentiable in S.

If $I(X)$ is $2p$-differentiable at a point X,

$$I(X)\otimes\nabla^2 G(X,\sigma)-\nabla^2 I(X)=\sum_k(2k)\sigma^{2(k-1)}\sum_{l,m,n}((2l)!!(2m)!!(2n)!!)^{-1}\times(\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n})+o(\sigma^{2(p-1)}),$$

where the sums run over $2\le k\le p$, l, m, $n\ge 0$, with $l+m+n=k$. Moreover, the relation holds uniformly in $X\in S$ if $I(X)$ is uniformly $2p$-differentiable in S.

Theorem 1.3.2.

If $I(X)$ is $2p$-differentiable at a point X,

$$2\sigma^{-2}[I(X)\otimes G(X,\sigma)-I(X)]-\nabla^2 I(X)=2\sum_k\sigma^{2(k-1)}\sum_{l,m,n}((2l)!!(2m)!!(2n)!!)^{-1}\times(\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n})+o(\sigma^{2(p-1)}),$$

where the sums run over $2\le k\le p$, l, m, $n\ge 0$, with $l+m+n=k$. This relation holds uniformly in $X\in S$ if $I(X)$ is uniformly 2p-differentiable in S. In particular, if $I(X)$ is 4-differentiable at a point X,

$$\lim_{\sigma\to 0}\sigma^{-2}[2\sigma^{-2}[I(X)\otimes G(X,\sigma)-I(X)]-\nabla^2 I(X)]=\nabla^4 I(X)/4,$$

and the relation holds uniformly in $X\in$ S if $I(X)$ is uniformly 4-differentiable in S.

Section 1.4 deals with the computation of gradients of arbitrary order.

Theorem 1.4.2.

Let $I(X)$ be a bounded continuous function, q-differentiable at a point X. If $Q(u,v,w)$ is any polynomial of total degree q in u, v, w, then with $Q(D)$ denoting the differential operator $Q(\partial/\partial x, \partial/\partial y, \partial/\partial z)$,

$$\lim_{\sigma\to 0} I(X) \otimes [Q(D)G(X,\sigma)] = Q(D)I(X),$$

the relation holding uniformly in $X \in S$ if $I(X)$ is uniformly q-differentiable in S. Moreover, if $I(X)$ is $2p+q$-differentiable at X, then

$$Q(D)[I(X)\otimes G(X,\sigma) - I(X)] = \sum_k \sigma^{2k} \sum_{\lambda,\mu,\nu} ((2\lambda)!!(2\mu)!!(2\nu)!!)^{-1} \times \left(\partial^{2k}[Q(D)I(X)]/\partial x^{2\lambda}\partial y^{2\mu}\partial z^{2\nu}\right) + o(\sigma^{2p}),$$

where the sums run over $2 \le k \le p$, $\lambda, \mu, \nu \ge 0$, with $\lambda+\mu+\nu = k$, and the relation holds uniformly in $X \in S$ if $I(X)$ is uniformly $2p+q$-differentiable in S.

In Section 1.5, the degree of approximation of C^2-functions by a Gaussian convolution is studied. The following are the main results:

Theorem 1.5.4.

Let $I(X) \in C$, and $\psi \in S$. Then:

(i) $I(X)\otimes G(X,\sigma) - I(X) = O(\psi(\sigma^2)), \sigma \to 0$, iff $K(I;t) = O(\psi(t^2))$, $t \to 0$;

(ii) $I(X)\otimes G(X,\sigma) - I(X) = o(\psi(\sigma^2)), \sigma \to 0$, iff $K(I;t) = o(\psi(t^2)), t \to 0$; and,

(iii) $I(X)\otimes G(X,\sigma) - I(X) \sim K(I;\sigma)$, if $K(I;t^{1/2}) \in S$.

Here, and in the sequel, S denotes the class of continuous non-decreasing functions $\psi(t)$ satisfying $0 = \psi(0) < \psi(t), t > 0$ such that $\lim_{\nu\to 0} \sup_{0<t\le 1/2} \nu\psi(t)/\psi(\nu t) = 0$.

Theorem 1.5.6.

Let $Q(D)I(X) \in C$, and $\psi \in S$. Then:

(i) $I(X)\otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = O(\psi(\sigma^2)), \sigma \to 0,$

iff $K(Q(D)I(X);t) = O(\psi(t^2)), t \to 0$;

(ii) $I(X)\otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = o(\psi(\sigma^2)), \sigma \to 0,$

iff $K(Q(D)I(X);t) = o(\psi(t^2)), t \to 0$; and,

(iii) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) \sim K(I;\sigma)$, if $K(Q(D)I(X);t^{1/2}) \in S$.

Here $Q(u,v,w)$ is a polynomial in u, v, w, as before, and $Q(D) \equiv Q(\partial/\partial x, \partial/\partial y, \partial/\partial z)$.

Saturation of convolutions with Gaussian filters has been discussed in section 1.6. Let the radial average $I_\rho(X)$ of a function $I(X) \in C$ be defined as:

$$I_\rho(X) = (4\pi\rho^2)^{-1} \iint_{A(X,\rho)} I(T)dA(T), \rho > 0,$$

where $A(X,\rho)$ denotes the surface of the sphere about the point X of radius ρ, T is a general point and $dA(T)$ stands for the area element on the surface about the point T. We call $I(X)$ harmonic if $I_\rho(X) = I(X)$, $X \in R^3$, $\rho > 0$. In particular, the following results are obtained:

Theorem 1.6.1.

Let $I(X) \in C$. Then:

(i) $|I(X) \otimes G(X,\sigma) - I(X)| = O(\sigma^2), \sigma \to 0$, iff $|I_r(X) - I(X)| = O(r^2), r \to 0$; and,

(ii) $|I(X) \otimes G(X,\sigma) - I(X)| = o(\sigma^2), \sigma \to 0$, iff $I(X)$ is harmonic.

Corollary 1.6.2.

Let $Q(D)I(X) \in C$. Then:

(i) $\|I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X)\| = O(\sigma^2), \sigma \to 0,$

iff $|Q(D)I_r(X) - Q(D)I(X)| = O(r^2), r \to 0$; and,

(ii) $|I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X)| = o(\sigma^2), \sigma \to 0,$

iff $Q(D)I(X)$ is harmonic.

In section 1.7, the following aysmptotic error estimates are obtained for 2-D Gaussian filter:

Theorem 1.7.1.

If $I(X)$ is 4-differentiable at a point X,

$$\lim_{\sigma\to 0}\sigma^{-2}\left\{I(X)\otimes\nabla^2 G(X,\sigma)-\nabla^2 I(X)\right\}=\nabla^4 I(X)/2,$$

and this relation holds uniformly in $X\in S\subset R^2$, if $I(X)$ is uniformly 4-differentiable in S. If $I(X)$ is $2p$-differentiable at a point X,

$$I(X)\otimes\nabla^2 G(X,\sigma)-\nabla^2 I(X)=\sum_k(2k)\sigma^{2(k-1)}\sum_{l,m}((2l)!!(2m)!!)^{-1}\times(\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m})+o(\sigma^{2(p-1)}),$$

where the sums run over $2\le k\le p$, l, $m\ge 0$, with $l+m=k$. Moreover, the relation holds uniformly in $X\in S$ if $I(X)$ is uniformly $2p$-differentiable in S.

Theorem 1.7.2.

If $I(X)$ is $2p$-differentiable at a point X,

$$[I(X)\otimes G(X,\sigma)-I(X)]=(\sigma^2/2)\nabla^2 I(X)+\sum_k\sigma^{2k}\sum_{l,m}((2l)!!(2m)!!)^{-1}(\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m})+o(\sigma^{2p}),$$

the sums being over $2\le k\le p$, with l, $m\ge 0$, and $l+m=k$. This relation holds uniformly in $X\in S\subset R^2$ if $I(X)$ is uniformly $2p$-differentiable in S. In particular, if $I(X)$ is 4-differentiable at a point X,

$$\lim_{\sigma\to 0}\sigma^{-2}\left\{2\sigma^{-2}[I(X)\otimes G(X,\sigma)-I(X)]-\nabla^2 I(X)\right\}=\nabla^4 I(X)/4,$$

and it holds uniformly in $X\in S$, if $I(X)$ is uniformly 4-differentiable in S.

Corollary 1.7.3.

For the difference of Gaussians (DOG) in two variables:

$$I(X)\otimes[G(X,\sigma_e)-G(X,\sigma_i)]-\sigma_e{}^2(1-r^2)\nabla^2 I(X)/2=\sigma_e{}^4(1-r^4)\nabla^4 I(X)/8+o(\sigma_e{}^4),\ \sigma_e\to 0.$$

The two-dimensional gradients of arbitrary order have been considered in section 1.8. The following is a pointwise simultaneous approximation property of an arbitrary order for the 2-D convolution.

Theorem 1.8.1.

Let $I(X)$ is a bounded continuous function, q-differentiable at a point X,

$$\lim_{\sigma\to 0} I(X)\otimes[Q(D)G(X,\sigma)] = Q(D)I(X),$$

the relation holding uniformly in $X \in S \subset \mathrm{R}^2$ if $I(X)$ is uniformly q-differentiable in S. Moreover, if $I(X)$ is $2p+q$-differentiable at X, then

$$Q(D)[I(X)\otimes G(X,\sigma) - I(X)] = \sum_k \sigma^{2k} \sum_{\lambda,\mu} ((2\lambda)!!(2\mu)!!)^{-1} \times \left(\partial^{2k}[Q(D)I(X)]/\partial x^{2\lambda}\partial y^{2\mu}\right) + o(\sigma^{2p}),$$

where the sums run over $2 \le k \le p$, $\lambda, \mu \ge 0$, with $\lambda + \mu = k$, and the relation holds uniformly in $X \in S \subset \mathrm{R}^2$, if $I(X)$ is uniformly $2p+q$-differentiable in S.

Error for smooth gradient functions in plane has been obtained in section 1.9 and the main results are as follows:

Theorem 1.9.4.

Let $I(X) \in C$, and $\psi \in S$. Then:

(i) $I(X)\otimes G(X,\sigma) - I(X) = O(\psi(\sigma^2)), \sigma \to 0$, iff $K(I;t) = O(\psi(t^2)), t \to 0$;

(ii) $I(X)\otimes G(X,\sigma) - I(X) = o(\psi(\sigma^2)), \sigma \to 0$, iff $K(I;t) = o(\psi(t^2)), t \to 0$; and,

(iii) $I(X)\otimes G(X,\sigma) - I(X) \sim K(I;\sigma)$, if $K(I;t^{1/2}) \in S$.

Theorem 1.9.6.

Let $Q(D)I(X) \in C$, and $\psi \in S$. Then:

(i) $I(X)\otimes[Q(D)G(X,\sigma)] - Q(D)I(X) = O(\psi(\sigma^2)), \sigma \to 0,$

iff $K(Q(D)I(X);t) = O(\psi(t^2)), t \to 0$;

(ii) $I(X)\otimes[Q(D)G(X,\sigma)] - Q(D)I(X) = o(\psi(\sigma^2)), \sigma \to 0,$

iff $K(Q(D)I(X);t) = o(\psi(t^2)), t \to 0$; and,

(iii) $I(X)\otimes[Q(D)G(X,\sigma)] - Q(D)I(X) \sim K(I;\sigma)$, if $K(Q(D)I(X);t^{1/2}) \in S$,

where $Q(u,v)$, as before, a polynomial in u, v, and $Q(D) \equiv Q(\partial/\partial x, \partial/\partial y)$.

Numerical simulations and discussions for the Gaussian filter have been done in section 1.10. Segmentation results with Gaussian filter corresponding to cases σ^{-1} = 64 and 32 are practically useful as the region of pathology is well segmented in these cases. Sobel maps for Gaussian filtered images give better description of the edges as compared to that of difference magnitudes of Gaussian filtered images. The contours obtained with Marr-Hildreth operator for σ = 2.0, 2.576 and 3.0 are also useful.

Chapter II is a study of Butterworth 1-D filters and its generalizations. These filters are characterized as solutions of certain minimization problems. The main results are:

Theorem 2.2.1.

For $f \in L_{2\pi}{}^2, m = 1,2,3,\ldots,$ and $\theta > 0$, the trigonometric polynomial $T_n(x) \in \mathcal{T}_n$ minimizing

$$Q_m = (2\pi)^{-1} \int_{[-\pi,\pi]} \left\{ f(x) - T_n(x)^2 + \theta n^{-m} T_n^{(m)}(x)^2 \right\} dx,$$

is unique and is given by

$$T_n(x) = T_n(f;x) = a_0(f)/2 + \sum_{1 \le k \le n} \left(a_k(f)\cos kx + b_k(f)\sin kx\right) / \left(1 + \theta (k/n)^{2m}\right),$$

where $a_k(f), b_k(f)$ are the Fourier coefficients of the function $f(x)$. The map $f \to T_n$ is linear, and, $T_n - f \to 0,\ n \to \infty$.

Here the norm $\|f\|$ is given by

$$\|f\| = \left(\int_{[-\pi,\pi]} |f(x)|^2 dx \right)^{1/2} = \pi \left\{ |a_0|^2/2 + \sum_{k \ge 1} \left(|a_k|^2 + |b_k|^2 \right) \right\}^{1/2}.$$

Theorem 2.2.3.

Let $f \in L_{2\pi}{}^2$, and $\psi \in S$. Then:

(i) There exists a constant $M \le \max(2,\theta)$, such that

$$\|f - T_n\| \le M K_{2m}(f; 1/n), n \to \infty;$$

(ii) $\|f - T_n\| = O(n^{-2m}), n \to \infty$, iff $f \in L_{2\pi,2m}{}^2$;

(iii) $f - T_n = o(n^{-2m}), n \to \infty$, iff f is a constant;

(iv) $f - T_n = O(\psi(n^{-2m})), n \to \infty$, iff $K_{2m}(f;t) = O(\psi(t^{2m})), t \to 0$;

(v) $|f - T_n| = o(\psi(n^{-2m})), n \to \infty$, iff $K_{2m}(f;t) = o(\psi(t^{2m})), t \to 0$; and,

(vi) $f - T_n \sim K_{2m}(f;1/n)$, if $K_{2m}(f;t^{1/(2m)}) \in S$.

Here $K_{2m}(f;t) = \inf_{g \in L_{2\pi,2m}{}^2} \{ |f-g| + t^{2m} |g|_{2m} \}$ $f \in L_{2\pi}{}^2$.

Section 2.3 studies the two dimensional Butterworth filter.

Theorem 2.3.1.

For $f \in L_{2\pi,2}{}^2$, $m = 1, 2, 3, \ldots$, and $\theta > 0$, the trigonometric polynomial $T_n(x,y) \in T_{n,2}$ minimizing

$$Q_m = (2\pi)^{-2} \int_{[-\pi,\pi]} \int_{[-\pi,\pi]} \{ |f(x,y) - T_n(x,y)|^2 + \theta \sum_{0 \le j \le m} {}^m C_j \, n^{-m} \partial^m T_n(x,y) / \partial x^{m-j} \partial y^j \, {}^2 \} dx dy,$$

is unique and is given by:

$$T_n(x,y) = T_n(f;x,y) = \sum_{1 \le |k| \le n} c_k(f) [1 + \theta(|k|/n)^{2m}]^{-1} \exp\{i(k_x x + k_y y)\},$$

where $c_k(f)$ are the complex Fourier coefficients of the function $f(x,y)$. The map: $f \to T_n$ is linear, and, $|T_n - f| \to 0$, $n \to \infty$.

The linear space invariant map $T_n : f \to T_n(f;x,y)$ has $|T_n| = 1$, and it can be evaluated by FFT using multiplicative replacements: $c_k \to c_k / [1 + \theta(|k|/n)^{2m}]$ in the frequency domain. The effect of this filtering for different θ and m and also the degree of approximation of $f(x,y)$ by the polynomials $T_n(f;x,y)$ is considered.

The main direct and inverse result on the degree of approximation of $f(x,y)$ by $T_n(x,y)$ is:

Theorem 2.3.3.

Let $f \in L_{2\pi,2}{}^2$, and $\psi \in S$. Then:

(i) There exists a constant $M \le \max(2,\theta)$, such that

$$\|f - T_n\| \le M K_{2m,2}(f; 1/n) \to 0, n \to \infty;$$

(ii) $\|f - T_n\| = O(n^{-2m}), n \to \infty$, iff $f \in L_{2\pi,2,2m}{}^2$;

(iii) $|f - T_n| = o(n^{-2m}), n \to \infty$, iff f is a constant;

(iv) $\|f - T_n\| = O(\psi(n^{-2m})), n \to \infty$, iff $K_{2m,2}(f;t) = O(\psi(t^{2m})), t \to 0$;

(v) $\|f - T_n\| = o(\psi(n^{-2m})), n \to \infty$, iff $K_{2m,2}(f;t) = o(\psi(t^{2m})), t \to 0$; and,

(vi) $\|f - T_n\| \sim K_{2m,2}(f; 1/n)$, if $K_{2m,2}(f; t^{1/(2m)}) \in S$.

Section 2.4 studies the three dimensional Butterworth filter, beginning with the following characterization:

Theorem 2.4.1.

For $f \in L_{2\pi,3}{}^2$, $m = 1, 2, 3, \ldots$, and $\theta > 0$, the trigonometric polynomial $T_n(x,y,z) \in T_{n,3}$ minimizing

$$Q_m = (2\pi)^{-3} \int_{[-\pi,\pi]\times[-\pi,\pi]\times[-\pi,\pi]} \Big\{ |f(x,y,z) - T_n(x,y,z)|^2 + \theta \sum_{0 \le i,j,l; i+j+l=m} [m!/(i!\,j!\,l!)] n^{-m} \partial^m T_n(x,y,z) / \partial x^i \partial y^j \partial z^l \Big\}^2 dx\,dy\,dz,$$

is unique and is given by:

$$T_n(x,y,z) = T_n(f;x,y,z) = \sum_{1 \le |k| \le n} c_k(f) \left[1 + \theta(|k|/n)^{2m}\right]^{-1} \exp\{i(k_x x + k_y y + k_z z)\},$$

where $c_k(f)$ are the complex Fourier coefficients of the function $f(x,y,z)$. The map: $f \to T_n$ is linear, and, $\|T_n - f\| \to 0$, $n \to \infty$.

As in the 2D case, the linear space invariant map $T_n : f \to T_n(f;x,y,z)$ has $\|T_n\| = 1$, and can be evaluated by FFT using multiplicative replacements: $c_k \to c_k / [1 + \theta(|k|/n)^{2m}]$ in the frequency domain. The main direct and inverse result on the degree of approximation of $f(x,y,z)$ by $T_n(x,y,z)$ is:

Theorem 2.4.3.

Let $f \in L_{2\pi,3}{}^2$, and $\psi \in S$. Then:

(i) There exists a constant $M \le \max(2, \theta)$, such that

$$\|f - T_n\| \le MK_{2m,3}(f;1/n) \to 0, n \to \infty;$$

(ii) $\|f - T_n\| = O(n^{-2m}), n \to \infty$, iff $f \in L_{2\pi,3,2m}{}^2$;

(iii) $\|f - T_n\| = o(n^{-2m}), n \to \infty$, iff f is a constant;

(iv) $\|f - T_n\| = O(\psi(n^{-2m})), n \to \infty$, iff $K_{2m,3}(f;t) = O(\psi(t^{2m})), t \to 0$;

(v) $\|f - T_n\| = o(\psi(n^{-2m})), n \to \infty$, iff $K_{2m,3}(f;t) = o(\psi(t^{2m})), t \to 0$; and,

(vi) $\|f - T_n\| \sim K_{2m,3}(f;1/n)$, if $K_{2m,3}(f;t^{1/(2m)}) \in S$.

Numerical results corresponding to 2-D Butterworth filter are discussed in section 2.5. Segmentations obtained with 2-D Butterworth filter for different values of parameters θ and m show similar results. Plots of percentage relative errors in the approximation with 2-D Butterworth for parameter θ with respect to different values of m show that the smoothing increases with the increase in the value of θ and decreases with the increase in the value of m.

Chapter III studies segmentations using the discrete convolutions with Korovkin, Jackson, generalized Jackson and De La Vallée-Poussin filters. Since their application to 2-D and 3-D images is iterative in nature, the degree of approximation results considered are 1-D. With the discretizations defined as:

$$K_{n,l(n)}(f;x) = [2/l(n)]\sum_{1\le k\le l(n)} f(t_{k,l(n)})\phi_n(t_{k,l(n)} - x),$$

where $\phi_n(x) = 1/2 + \sum_{1\le k\le n} \rho_{k,l(n)} \cos kx \ge 0$, $x \in [0,2\pi]$.

For a general $K_{n,l(n)}$, following is the asymptotic error formula in the approximation of $f(x)$:

Theorem 3.2.2.

If $l(n) > m(n) + 2$, f is a bounded 2π-periodic function, and f'' exists at a point x,

$$K_{n,l(n)}(f;x) - f(x) = (1 - \rho_{1,n})f''(x) + o(1 - \rho_{1,n}),$$

iff for some $k \ge 2$,

$$\lim_{n\to\infty} (1 - \rho_{k,n})/(1 - \rho_{1,n}) = k^2.$$

Section 3.3 studies the direct and inverse results on the approximation of Korovkin, Jackson and generalized Jackson operators. Let $C_{2\pi}$ as usual denote the space of 2π-periodic continuous functions on R, and $C_{2\pi}{}^2$ the sub-space of functions having continuous second order derivative. Let $\|f\| = \max |f(x)|$, $f \in C_{2\pi}$, and let $\|g\|_2 = \|g\| + \|g''\|$, $g \in C_{2\pi}{}^2$. Let the Peetre's K functional for $f \in C_{2\pi}$ be defined as:

$$K(f;t) = \inf\{\|f-g\| + t^2 \|g''\| : g \in C_{2\pi}{}^2\}, t > 0.$$

The direct and inverse results on the degree of approximation of f by the discretizations $K_{n,l(n)}$ are as follows:

Theorem 3.3.3.

Let $f(x) \in C_{2\pi}$, and $\psi \in S$. If $1-\rho_{1,n} = O([m(n)]^{-2}) \sim 1/n^2$ and $l(n) \geq m(n)+2$, then:

(i) $\|K_{n,l(n)}f - f\| \leq AK(f;(1-\rho_{1,n})^{1/2}) \to 0, n \to \infty$, where A is a constant;

(ii) $\|K_{n,l(n)}f - f\| = O(\psi(n^{-2})), n \to \infty$, iff $K(f;t) = O(\psi(t^2)), t \to 0$;

(iii) $\|K_{n,l(n)}f - f\| = o(\psi(n^{-2})), n \to \infty$, iff $K(f;t) = o(\psi(t^2)), t \to 0$;

(iv) $\|K_{n,l(n)}f - f\| \sim K(f;n^{-1})$, if $K(f;t^{1/2}) \in S$.

For the De La Vallée-Poussin kernel there holds: $1-\rho_{1,n} = O(1/n)$. Hence the above results are not applicable to the discrete filters $V_{n,l(n)}(f;x)$, which are studied separately in section 3.4.

However, since the relation $1-\rho_{1,n} = O([m(n)]^{-2}) \sim 1/n^2$ is valid for the earlier mentioned Korovkin, the generalized Jackson, and Jackson kernels (the case $p = 2$, in the following), for the discrete convolutions $A_{n-2,l(n)}(f;x)$, $L_{np-p,l(n)}(f;x)$ and $L_{2n-2,l(n)}(f;x)$, we obtain the following:

Corollary 3.3.5.

Let $f(x) \in C_{2\pi}$, and $\psi \in S$. Then, if $l(n) \geq n$:

(i) $A_{n-2,l(n)}f - f \le A\omega_2(f;1/n) \to 0, n \to \infty$, for a constant A;

(ii) $A_{n-2,l(n)}f - f = O(\psi(n^{-2})), n \to \infty$, iff $\omega_2(f;t) = O(\psi(t^2)), t \to 0$;

(iii) $A_{n-2,l(n)}f - f = o(\psi(n^{-2})), n \to \infty$, iff $\omega_2(f;t) = o(\psi(t^2)), t \to 0$;

(iv) $A_{n-2,l(n)}^{t}f - f \sim \omega_2(f;n^{-1})$, if $\omega_2(f;t^{1/2}) \in S$.

Corollary 3.3.6.

Let $f(x) \in C_{2\pi}$, $p \ge 2$, and $\psi \in S$. Then, if $l(n) \ge np - p + 2$:

(i) $L_{np-p,l(n)}f - f \le A\omega_2(f;1/n) \to 0, n \to \infty$, for a constant A;

(ii) $L_{np-p,l(n)}f - f = O(\psi(n^{-2})), n \to \infty$, iff $\omega_2(f;t) = O(\psi(t^2)), t \to 0$;

(iii) $L_{np-p,l(n)}f - f = o(\psi(n^{-2})), n \to \infty$, iff $\omega_2(f;t) = o(\psi(t^2)), t \to 0$;

(iv) $L_{np-p,l(n)}f - f \sim \omega_2(f;n^{-1})$, if $\omega_2(f;t^{1/2}) \in S$.

In particular for the Bernsteinian orders $n^{-\alpha}$, $0 < \alpha < 2$, there hold:

Corollary 3.3.7.

Let $f(x) \in C_{2\pi}$, $0 < \alpha < 2$, and $l(n) \ge n$. Then:

(i) $A_{n-2,l(n)}f - f = O(n^{-\alpha}), n \to \infty$, iff $\omega_2(f;t) = O(t^{\alpha}), t \to 0$;

(ii) $A_{n-2,l(n)}f - f = o(n^{-\alpha}), n \to \infty$, iff $\omega_2(f;t) = o(t^{\alpha}), t \to 0$;

(iii) $A_{n-2,l(n)}f - f \sim n^{-\alpha}$, if $\omega_2(f;t) \sim t^{\alpha}$.

Corollary 3.3.8.

Let $f(x) \in C_{2\pi}$, $0 < \alpha < 2$, $p \ge 2$, and $l(n) \ge np - p + 2$. Then:

(i) $L_{np-p,l(n)}f - f = O(n^{-\alpha}), n \to \infty$, iff $\omega_2(f;t) = O(t^{\alpha}), t \to 0$;

(ii) $L_{np-p,l(n)}f - f = o(n^{-\alpha}), n \to \infty$, iff $\omega_2(f;t) = o(t^{\alpha}), t \to 0$;

(iii) $L_{np-p,l(n)}f - f \sim n^{-\alpha}$, if $\omega_2(f;t) \sim t^{\alpha}$.

Approximation of De La Vallée-Poussin is studied in Section 3.4, where the following results are proved:

Theorem 3.4.5. (Approximation of derivatives).

If f is a bounded 2π-periodic function such that for some natural number m, and at some point x, $f^{(m)}(x)$ exists, then with $l(n) > n+[m/2]+1$:

$$\lim_{n\to\infty} V_{n,l(n)}{}^{(m)}(f;x) = f^{(m)}(x).$$

Moreover, if $f^{(m)}(x)$ exists for all x and belongs to $C_{2\pi}$, the convergence is uniform in x.

Theorem 3.4.6. (Voronovskaya Type Asymptotic Formula).

If f is a bounded 2π-periodic function such that for a natural number m, at some point x, $f^{(m+2)}(x)$ exists, then with $l(n) > n+[m/2]+2$:

$$\lim_{n\to\infty} n\left(V_{n,l(n)}{}^{(m)}(f;x) - f^{(m)}(x)\right) = f^{(m+2)}(x).$$

Moreover, if $f^{(m+2)}(x)$ exists for all x and belongs to $C_{2\pi}$, the convergence is uniform in x.

Let $C_{2\pi}{}^{m}$, $(m \geq 0)$, denote the space of 2π-periodic functions having continuous m-th order derivative on the real line. Let $\|f\| = \max |f(x)|, f \in C_{2\pi}$. Let a Peetre's K-functional, related with the pair $(C_{2\pi}{}^{m}, C_{2\pi}{}^{m+2})$, for $f \in C_{2\pi}{}^{m}$, be defined as:

$$K_{m,m+2}(f;t) = \inf\left\{ \|f^{(m)} - g^{(m)}\| + t^2 \|g^{(m+2)}\| : g \in C_{2\pi}{}^{m+2} \right\}, t > 0.$$

In order to handle the approximation of functions as well as their derivatives simultaneously, we would naturally identify $f^{(0)}(x)$ with $f(x)$, and, $C_{2\pi}{}^{0}$ with $C_{2\pi}$.

The following is the main direct and inverse result about approximation of f by $V_{n,l(n)}f$ (case $m = 0$) and its continuous derivatives $f^{(m)}$ by the m-th derivative $V_{n,l(n)}{}^{(m)}f$ of $V_{n,l(n)}f$:

Theorem 3.4.10.

Let $f(x) \in C_{2\pi}{}^{m}$, $m \geq 0$, $l(n) > n+[m/2]+2$, and $\psi \in S$. Then:

(i) $V_{n,l(n)}{}^{(m)} f - f^{(m)} \leq AK_{m,m+2}(f, n^{-1/2}) \to 0, n \to \infty$, A being a constant;

(ii) $V_{n,l(n)}{}^{(m)} f - f^{(m)} = O(\psi(n^{-1})), n \to \infty$, iff $K_{m,m+2}(f;t) = O(\psi(t^2)), t \to 0$;

(iii) $V_{n,l(n)}{}^{(m)} f - f^{(m)} = o(\psi(n^{-1})), n \to \infty$, iff $K_{m,m+2}(f;t) = o(\psi(t^2)), t \to 0$;

(iv) $V_{n,l(n)}{}^{(m)} f - f^{(m)} \sim K_{m,m+2}(f, n^{-1/2})$, if $K_{m,m+2}(f;t^{1/2}) \in S$.

Numerical results on the approximation of Korovkin, Jackson, generalized Jackson and De La Vallée-Poussin are presented in sections 3.5 and 3.6. Blurring is more on images convolved with Korovkin filter as compared to with that of Jackson and Generalized Jackson operators corresponding to the cases n = 128, 64, 32, 16 and 8. Segmentation result obtained with Korovkin operator for n = 128 seems to be practically useful. Spatial convolutions with De La Vallée-Poussin operator for n = 128, 64, 32, 16 and 8 gives too much blurring, but the result obtained corresponding to case n = 4096 is useful for practical purposes.

In **Chapter IV**, the segmentation methods based on maximum likelihood estimation in univariate and multivariate cases are discussed. The segmentation algorithm for the single contrast is given in section 4.2. Section 4.3 contains the derivation of iterative scheme and the corresponding algorithm for multivariate segmentation (R.K.S. Rathore, S. Datta, R.K. Gupta, S.B. Rao and R. Verma [2001]). A detailed study of two pathologies - presence of edema in brain parenchyma, and post-traumatic epilepsy with perilesional gliosis, using segmentation methods discussed in previous sections, has been done in sections 4.4 and 4.5. Tracing of boundaries using cubic splines (S. Datta, R.K. Gupta, S.B. Rao, V.S.N. Kaliprasad and R.K.S. Rathore [2000]) is described in section 4.6.

CHAPTER I : THE GAUSSIAN FILTER AND THE MARR-HILDRETH AND RELATED APPROXIMATIONS

This chapter studies the degree of approximation properties of the Marr-Hildreth and other related convolutions with the Gaussian function and the resulting segmentations. Convolution of an image with Gaussian filter is described in introductory section. Section 1.2 contains the basic results on degree of approximation by a sequence of linear operators which are required for the subsequent analysis. Section 1.3 contains the main result of asymptotic estimates of Marr-Hildreth operator. Computation of gradients of arbitrary order has been done in section 1.4 and the degree of approximation by Gaussian convolution has been studied in section 1.5. Section 1.6 discusses the saturation of convolutions with Gaussian filters and the asymptotic error estimates for 2-D Gaussian filters are obtained in section 1.7. 2-D gradients of arbitrary order are considered in section 1.8 and the error for smooth gradient functions in plane is studied in section 1.9. Finally, section 1.10 includes the numerical simulations and discussions for the Gaussian filter and Marr-Hildreth operator.

1.1. INTRODUCTION

MRI produces 2-D cross-sectional slices containing the 3-D information of the examined organs. In the commonly used method where the physician analyzes the sequence slice by slice, it is very difficult to reconstruct mentally the 3-D shape of the structures explored. To overcome this problem by computer-generated perspective 3-D displays of surfaces, the surfaces of the objects to be displayed must be determined. Typically, this is done manually slice by slice. Due to their characteristic gray levels, bony structures in CT images are amenable to threshold methods. Extraction of the surface of soft tissue objects like brain or muscles is much more difficult, as the

contrast between these objects is poor. In MR images, even though the contrast between different tissues is much greater than in CT images, there is not a unique correspondence of gray level ranges to different tissues. This necessitates the use of an edge detection operator to find the surfaces in the image sequence allowing a segmentation of main anatomical constituents an anatomist would like to dissect. As discontinuities lead to wrong or misleading structures in the images of the objects, for a 3-D reconstruction, it is very important that the surfaces found are closed in 3-D space. This requires a segmentation method that recognizes the information in the data volume in all three dimensions.

Following M. Bomans, K-H. Höhne, U. Tiede, and M. Riemer [1990], among various algorithms for edge detection, the Marr-Hildreth operator and related approximations: (i) easily extend to 3-D and deliver closed contours without much computational effort, (ii) calculate the surfaces based on larger neighborhoods than classical edge detection operators, and (iii) work, as claimed in D. Marr and E.C. Hildreth [1980], similarly to the low-level processing of the human visual system. The 2-D Marr-Hildreth operator is defined (E.C. Hildreth [1983]) as:

$$C(x,y) = \nabla^2 (I(x,y) \otimes G(x,y,\sigma)),$$

where $I(x,y)$ is an image,

$$G(x,y,\sigma) = (2\pi\sigma^2)^{-1} \exp[-(x^2+y^2)/(2\sigma^2)],$$

is the Gaussian function,

$$\nabla^2 \equiv (\partial^2/\partial x^2 + \partial^2/\partial y^2),$$

is the Laplacian, and $C(x,y)$ is the contour image. The convolution ($\otimes$) in 2-D is defined as:

$$I(x,y) \otimes G(x,y,\sigma) = \int_{(-\infty,\infty)}\int_{(-\infty,\infty)} I(t,s) * G(x-t, y-s, \sigma) dtds.$$

The Marr-Hildreth operator first smoothens the data volume to remove structures of higher frequency components, thereby shortening the frequency range on which intensity changes can occur, and only intensity changes at a specific resolution are considered. Marr and Hildreth advocate Gaussian operator as an optimal compromise between conflicting smoothness constraints as Gaussian distribution is

smooth and localized both in the spatial and frequency domain. It is thus unlikely that artifacts are introduced by the filtering process. Another advantage of the Gaussian function is that the amount of smoothing and thereby the considered frequency range can easily be adjusted by choosing an appropriate value for the standard deviation. If a low standard deviation is taken, only values inside a small neighborhood are smoothed; with a greater standard deviation, values over a larger neighborhood are smoothed.

If there is an intensity change with a particular orientation in an image, there will usually be a peak in the first, and a zero crossing in the second directional derivative taken perpendicular to the orientation of the intensity change. The detection of edges with a directional operator, however, has the disadvantage that the operator must be applied at different directions in order to get intensity changes of all orientations. Marr and Hildreth chose the Laplacian as a rotationally invariant operator, which therefore detects edges of any orientation. The 3-D extension of the Marr-Hildreth operator is defined by

$$C(x,y,z) = \nabla^2 (I(x,y,z) \otimes G(x,y,z,\sigma)) = I(x,y,z) \otimes \nabla^2 G(x,y,z,\sigma),$$

$$G(x,y,z,\sigma) = (2\pi\sigma^2)^{-3/2} \exp[-(x^2+y^2+z^2)/(2\sigma^2)],$$

$$\nabla^2 \equiv (\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2),$$

$$\nabla^2 G(x,y,z,\sigma) = (2\pi)^{-3/2} \exp[-(x^2+y^2+z^2)/(2\sigma^2)][(x^2+y^2+z^2)/\sigma^2 - 3]/\sigma^5.$$

As the contour volume can be calculated by convolving the data volume with the Laplacian of a Gaussian function, the operator $\nabla^2 G(x,y,z,\sigma)$ is called "Laplacian of a Gaussian" or $\nabla^2 G$ -operator.

The 3-D convolution with $\nabla^2 G$ is not separable into three one-dimentional convolutions. Nevertheless, the 2-D Laplacian is decomposable into two separable functions (A. Huertas and G. Medioni [1986]), which for 3-D, results in the following:

$$\nabla^2 G(x,y,z,\sigma) = G''(x,\sigma) * G(y,\sigma) * G(z,\sigma) + G(x,\sigma) * G''(y,\sigma) * G(z,\sigma) + G(x,\sigma) * G(y,\sigma) * G''(z,\sigma),$$

where

$$G(\xi,\sigma) = (2\pi\sigma^2)^{-1/2} \exp[-(\xi^2)/(2\sigma^2)],$$

$$G''(\xi,\sigma) = (2\pi\sigma^2)^{-1/2} (\xi^2/\sigma^2 - 1) \exp[-(\xi^2)/(2\sigma^2)].$$

The algorithm needs eight 1-D convolutions and a memory of 3.X.Y.M elements (X,Y: size of the data cube; M: filter width) for obtaining the convolution. Also Marr and Hildreth have shown that their operator can be approximated by a separable operator, the "Difference of Gaussians" (DOG) operator:

$$\begin{aligned}\nabla^2 G(x,y,z,\sigma) &\approx G(x,y,z,\sigma_e) - G(x,y,z,\sigma_i) \\ &\approx G(x,\sigma_e) * G(y,\sigma_e) * G(z,\sigma_e) - G(x,\sigma_i) * G(y,\sigma_i) * G(z,\sigma_i),\end{aligned}$$

which needs only six 1-D convolutions and approximately 2.X.Y.M memory elements. For implementation it has been suggested to prefer the DOG operator for its lower memory requirements and smaller number of convolutions (M. Bomans, K-H. Höhne, U. Tiede and M. Riemer [1990]).

It is reported that if the DOG operator is used as an approximation for the $\nabla^2 G$ operator, σ_e and σ_i have to be determined such that the bandwidth of the filter is small and the sensitivity is large: if $\sigma_i : \sigma_e$ approximates 1, the DOG and the $\nabla^2 G$ operator are (up to a constant factor) identical, but the sensitivity is 0, and if the ratio increases, the sensitivity increases, but the approximation becomes worse and the bandwidth of the filter enlarges. As a compromise, Marr and Hildreth chose a value of 1.6. Some mathematical properties of the Marr-Hildreth operator are studied by V. Torre and T. Poggio [1986]. Reportedly the Marr-Hildreth operator detects coarse outline of different tissues e.g., skin or ventricular system quite well, but in the region of brain sulci the contours found relate to white-gray matter transition than to the transition from gray matter to CSF. To overcome the problem it is suggested to shift the contours in the sulci outwards using morphological dilation, erosion and closing filters for two dimensional objects that alter the form of objects and recognize special objects (R.M. Haralick, S.R. Sternberg and X. Zhuang [1987], J. Serra [1982]).

In view of the fact that it is not very clear as to precisely what smoothing and filtering does to the contours, edges and the boundaries associated with different segments or objects, we look at the effect of the degree of approximation on the segmentations and associated curves. The experiments in subsequent chapters seem to

confirm that at the present de-facto standard of 256×256 resolution, the segmentation properties depend essentially on the degree of approximation of noise reducing basic operators which are essentially convolutions based bell shaped functions.

1.2. BASIC LEMMAS ON DEGREE OF APPROXIMATION

Some basic results from R.K.S. Rathore [2000], for studying degree of approximation by a sequence of linear operators $\{L_n\}$, are summarized as follows: Let $Y \subset X$ be Banach spaces and Φ a semi-norm for which $Y = \{f \in X : \Phi(f) < \infty\}$, and let

$$K(f;t) = \inf_{g \in Y} \{ \| f - g \|_X + t\Phi(g) \}.$$

Such a K is called a Peetre's K-functional corresponding to the triple $\{X, Y, \Phi\}$.

For a sequence $\{L_n\}$ of linear operators from $X \to Y$, and a sequence $\{\sigma_n\}$ of positive numbers such that $\sigma_{n+1} \geq c\sigma_n \to 0, n \to \infty$, an inequality of the type

$$\| L_n f - f \|_X \leq AK(f;\sigma_n),$$

is called a Jackson inequality,

$$\Phi(L_n f) \leq B\sigma_n^{-1} \| f \|_X,$$

a Bernstein inequality, and

$$\Phi(L_n g) \leq C\Phi(g), g \in Y,$$

the uniform Φ-boundedness of $\{L_n\}$ over Y. In such a case we say that $\{Y, \Phi, K\}$ is a J-B triple with the associated rate sequence $\{\sigma_n\}$.

Lemma 1.2.1.

For a continuous non-decreasing function $\psi(t)$ satisfying $0 = \psi(0) < \psi(t)$ for $t > 0$, the following statements are equivalent:

(i) $\int_{[\delta,1]} [\psi(t)/t^2] dt = O(\psi(\delta))$;

(ii) $\int_{[\delta,1]} [\psi(t)/t^2] dt \sim \psi(\delta)$;

(iii) For some $c>1$, $\limsup_{t\to 0}\psi(ct)/\psi(t)<c$;

(iv) For some $A_0>0$, $\psi(A\delta)/\psi(\delta)\le(1/2)A, (A\ge A_0, A\delta\le 1/2)$;

(v) $\lim_{v\to 0}\sup_{0<t\le 1/2} v\psi(t)/\psi(vt)=0$;

(vi) $\sup_{0<t\le 1/2} v\psi(t)/\psi(vt)=O(1/|\ln v|)\to 0, v\to 0$.

If ψ is as in Lemma 1.2.1, and satisfies any of the equivalent conditions (i)-(vi) in it, we say that $\psi\in S$ (the class of steady order functions).

Lemma 1.2.2.

Let $\{\sigma_n\}$ satisfy $\sigma_{n+1}\ge c\sigma_n\to 0, n\to\infty$, $K(t)\ge 0$ be non-decreasing, and $\delta(t)\ge 0, t\in(0,1]$. Let $K(\sigma_j)\le M\{\delta(\sigma_n)+\sigma_j K(\sigma_n)/\sigma_n\}$, $1\le n<j<\infty$, for a constant M. If $\psi\in S$, then:

(i) $K(t)=O(\psi(t))$, if $\delta(\sigma_n)=O(\psi(\sigma_n))$;

(ii) $K(t)=o(\psi(t))$, if $\delta(\sigma_n)=o(\psi(\sigma_n))$; and,

(iii) $K(\sigma_n)=O(\delta(\sigma_n))$, if $K\in S$.

1.3. ASYMPTOTIC ESTIMATES FOR MARR-HILDRETH OPERATOR

Several authors, including K. Yoshida [1971], P.L. Butzer and H. Berens [1967], H.S. Shapiro [1969], etc. have studied approximation properties of one dimensional convolution with the Gaussian filter:

$$G(x,\sigma)=(2\pi\sigma^2)^{-1/2}\exp[-x^2/(2\sigma^2)].$$

These studies enable a better understanding of the Marr-Hildreth operator. In what follows, we consider the pointwise approximation of gradients of a gerneral order by the corresponding gradients of the 3-D Gaussian filter. The main results clarify the asymptotic dependence of the degree of approximation on the gradients of higher orders as well as the interplay of the smoothness of the gradient in its degree of

approximation. We begin by developing an asymptotic error estimate and a more detailed asymptotic expansion of it for the Marr-Hildreth operator.

A bounded continuous function $f(X) \equiv f((x, y, z)')$ is called m-differentiable at a point $X \in \mathrm{R}^3$, if it admits the following expansion in a neighborhood of the point X:

$$f(X+H) = f(X) + \sum_{1 \le k \le m} (k!)^{-1} (H \cdot \nabla)^k f(X) + \mathrm{o}(H^{\,m}),\ H \to 0.$$

The function $f(X)$ is called uniformly m-differentiable on a set $S \subset \mathrm{R}^3$, if given an arbitrary $\varepsilon > 0$, there exists a $\delta > 0$, independent of X, such that whenever $H < \delta$,

$$f(X+H) - \left\{ f(X) + \sum_{1 \le k \le m} (k!)^{-1} (H \cdot \nabla)^k f(X) \right\} < \varepsilon H^{\,m},\ X \in \mathrm{S}.$$

In conformity with $\nabla^2 \equiv (\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)$, let

$$\nabla^4 \equiv (\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)^2.$$

In the sequel $n!!$, the semi-factorial of n, is the product of all even or odd natural numbers, according as n is even or odd. For the 3-D Gaussian filter

$$G(X, \sigma) = (2\pi\sigma^2)^{-3/2} \exp[-X^{\,2}/(2\sigma^2)],$$

there holds the following asymptotic error estimate:

Theorem 1.3.1.

If $I(X)$ is 4-differentiable at a point X,

$$\lim_{\sigma \to 0} \sigma^{-2} \left\{ I(X) \otimes \nabla^2 G(X, \sigma) - \nabla^2 I(X) \right\} = \nabla^4 I(X)/2,$$

and this relation holds uniformly in $X \in S$ if $I(X)$ is uniformly 4-differentiable in S.

If $I(X)$ is $2p$-differentiable at a point X,

$$I(X) \otimes \nabla^2 G(X, \sigma) - \nabla^2 I(X) = \sum_k (2k) \sigma^{2(k-1)} \sum_{l,m,n} ((2l)!!(2m)!!(2n)!!)^{-1} \times (\partial^{2k} I(X)/\partial x^{2l} \partial y^{2m} \partial z^{2n}) + \mathrm{o}(\sigma^{2(p-1)}),$$

where the sums run over $2 \le k \le p$, l, m, $n \ge 0$, with $l+m+n = k$. Moreover, the relation holds uniformly in $X \in S$ if $I(X)$ is uniformly $2p$-differentiable in S.

Proof: In the following a triple integral $\int\int\int_{\cdots}[.]dH$ is an integral over R^3, where in each of the iterated integrals the limits are from $-\infty$ to $+\infty$. We have

$$I(X)\otimes\nabla^2 G(X;\sigma) = I(X)\otimes\left((2\pi)^{-3/2}\sigma^{-5}\exp\left\{-|X|^2/(2\sigma^2)\right\}(-3+|X|^2/\sigma^2)\right)$$

$$= \int\int\int_{\cdots}[I(X)+\sum_{1\le k\le 2p}(k!)^{-1}(-H\cdot\nabla)^k I(X)+o(|H^{2p})]$$

$$\times(2\pi)^{-3/2}\sigma^{-5}\exp\left\{-H^2/(2\sigma^2)\right\}(-3+H|^2/\sigma^2)dH.$$

Using $(2\pi)^{-1/2}\sigma^{-1}\int_{\cdot} t^j\exp(-t^2/2\sigma^2)dt = 0$, if j is odd and $(j\text{-}1)!!\sigma^j$, if j is even

$$\int\int\int_{\cdots}\left(\sum_{0\le k\le 2p}(k!)^{-1}(-H\cdot\nabla)^k I(X)\right)(2\pi)^{-3/2}\sigma^{-5}\exp\left\{-|H|^2/(2\sigma^2)\right\}(-3+|H|^2/\sigma^2)dH$$

$$= \sum_{0\le k\le p}((2k)!)^{-1}\sum_{l,m,n\ge 0,l+m+n=2k}[(2k)!/(l!m!n!)](\partial^{2k}I(X)/\partial x^l\partial y^m\partial z^n)(2\pi)^{-3/2}\sigma^{-5}$$

$$\times\int\int\int_{\cdots}\lambda^l\mu^m\nu^n\exp\{-(\lambda^2+\mu^2+\nu^2)/(2\sigma^2)\}[-3+(\lambda^2+\mu^2+\nu^2)/\sigma^2]d\lambda d\mu d\nu$$

$$= \sum_{0\le k\le p}\sum_{l,m,n\ge 0,l+m+n=k}[1/((2l)!(2m)!(2n)!)](\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n})(2\pi)^{-3/2}\sigma^{-5}$$

$$\times\int\int\int_{\cdots}\lambda^{2l}\mu^{2m}\nu^{2n}\exp\{-(\lambda^2+\mu^2+\nu^2)/(2\sigma^2)\}[-3+(\lambda^2+\mu^2+\nu^2)/\sigma^2]d\lambda d\mu d\nu$$

$$= \sigma^{-2}\sum_{0\le k\le p}\sum_{l,m,n\ge 0,l+m+n=k}[1/((2l)!(2m)!(2n)!)](\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n})$$

$$\times[(2l-1)!!(2m-1)!!(2n-1)!!(-3+(2l+1+2m+1+2n+1))]\sigma^{(2l+2m+2n)}$$

$$= \sum_{1\le k\le p}\sum_{l,m,n\ge 0,l+m+n=k}[2k/((2l)!!(2m)!!(2n)!!)]\sigma^{2(l+m+n-1)}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n}$$

$$= (\partial^2/\partial x^2+\partial^2/\partial y^2+\partial^2/\partial z^2)I(X)+\sum_{2\le k\le p}\sum_{l,m,n\ge 0,l+m+n=k}2k/((2l)!!(2m)!!(2n)!!)$$

$$\times\sigma^{2(k-1)}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n}.$$

For the $o(H^{2p})$-term, given an arbitrary $\varepsilon > 0$, there exists a $\delta > 0$, such that $o(H^{2p}) < (\varepsilon + H^2/\delta^2)H^{2p}$, $H\in R^3$. Using,

$$\int\int\int_{\cdots} X^{2p+2j}G(X,\sigma)dX = O(\sigma^{2p+2j}),\sigma\to 0,$$

the net contribution of the $o(H^{2p})$–term being $o(\sigma^{2(p-1)})$, and the result for $2p$-differentiable $I(X)$ follows. In particular, for $p = 2$,

$\lim_{\sigma\to 0}\sigma^{-2}[I(X)\otimes\nabla^2 G(X;\sigma)-\nabla^2 I(X)]$

$$= 4\sum_{l,m,n\geq 0, l+m+n=2} ((2l)!!(2m)!!(2n)!!)^{-1} \partial^{2k} I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n}$$

$$= 4[\{\partial^4/\partial x^4 + \partial^4/\partial y^4 + \partial^4/\partial z^4\} I(X)/8 + \{\partial^4/\partial y^2\partial z^2 + \partial^4/\partial z^2\partial x^2 + \partial^4/\partial x^2\partial y^2\} I(X)/4]$$

$$= (1/2)(\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)^2 I(X).$$

The uniformity of the asymptotic estimates in the cases of uniform differentiability is clear as the δ in above could be chosen to be independent of $X \in S$.

In the sequel, a similar order of approximation error would be observed for a large class of 3-D convolutions with combinations of the Gaussian and other filter functions, which at the same time are separable into three 1-D convolutions. In particular, for the error in the pointwise approximation by the Gaussian filter:

Theorem 1.3.2.

If $I(X)$ is $2p$-differentiable at a point X,

$$2\sigma^{-2}[I(X)\otimes G(X,\sigma) - I(X)] - \nabla^2 I(X) = 2\sum_k \sigma^{2(k-1)} \sum_{l,m,n} ((2l)!!(2m)!!(2n)!!)^{-1} \times (\partial^{2k} I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n}) + o(\sigma^{2(p-1)}),$$

where the sums run over $2 \leq k \leq p$, l, m, $n \geq 0$, with $l+m+n = k$. This relation holds uniformly in $X \in S$ if $I(X)$ is uniformly 2p-differentiable in S. In particular, if $I(X)$ is 4-differentiable at a point X,

$$\lim_{\sigma\to 0} \sigma^{-2}[2\sigma^{-2}\{I(X)\otimes G(X,\sigma) - I(X)\} - \nabla^2 I(X)] = \nabla^4 I(X)/4,$$

and the relation holds uniformly in $X \in S$ if $I(X)$ is uniformly 4-differentiable in S.

Proof: Following the proof of the previous theorem:

$$I(X)\otimes G(X,\sigma) = I(X)\otimes\left((2\pi)^{-3/2}\sigma^{-3}\exp\left\{-\|X\|^2/(2\sigma^2)\right\}\right)$$

$$= \int_{\cdots}^{\cdots}\left\{I(X) + \sum_{1\leq k\leq 2p}(k!)^{-1}(-H\cdot\nabla)^k I(X) + o(\|H\|^{2p})\right\}(2\pi)^{-3/2}\sigma^{-3} \times \exp\left\{-\|H\|^2/(2\sigma^2)\right\} dH.$$

As before, the $o(\|H\|^{2p})$-term on the right hand side contributes to $o(\sigma^{2p})$, which holds uniformly in $X \in S$, in the uniform differentiability case. For the remaining term

$$\int\!\!\int\!\!\int \Big(\sum_{0\le k\le 2p}(k!)^{-1}(-H\cdot\nabla)^k I(X)\Big)(2\pi)^{-3/2}\sigma^{-3}\exp\Big\{-X^2/(2\sigma^2)\Big\}dH$$

$$=\sum_{0\le k\le p}((2k)!)^{-1}\sum_{l,m,n\ge 0,l+m+n=2k}[(2k)!/(l!m!n!)](\partial^{2k}I(X)/\partial x^l\partial y^m\partial z^n)(2\pi)^{-3/2}\sigma^{-3}$$

$$\times\int\!\!\int\!\!\int \lambda^l\mu^m\nu^n\exp\{-(\lambda^2+\mu^2+\nu^2)/(2\sigma^2)\}d\lambda d\mu d\nu$$

$$=\sum_{0\le k\le p}\sum_{l,m,n\ge 0,l+m+n=k}[1/((2l)!(2m)!(2n)!)](\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n})(2\pi)^{-3/2}\sigma^{-3}$$

$$\times\int\!\!\int\!\!\int \lambda^{2l}\mu^{2m}\nu^{2n}\exp\{-(\lambda^2+\mu^2+\nu^2)/(2\sigma^2)\}d\lambda d\mu d\nu$$

$$=\sum_{0\le k\le p}\sum_{l,m,n\ge 0,l+m+n=k}[(2l)!(2m)!(2n)!]^{-1}(\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n})$$

$$\times[(2l-1)!!(2m-1)!!(2n-1)!!]\sigma^{(2l+2m+2n)}$$

$$=\sum_{0\le k\le p}\sum_{l,m,n\ge 0,l+m+n=k}((2l)!!(2m)!!(2n)!!)^{-1}\sigma^{2k}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n}.$$

Hence,

$$I(X)\otimes G(X,\sigma)=I(X)+\sigma^2(\partial^2/\partial x^2+\partial^2/\partial y^2+\partial^2/\partial z^2)I(X)/2$$
$$+\sigma^4(\partial^2/\partial x^2+\partial^2/\partial y^2+\partial^2/\partial z^2)^2 I(X)/8$$
$$+\sum_{3\le k\le p}\sigma^{2k}\sum_{l,m,n\ge 0,l+m+n=k}((2l)!!(2m)!!(2n)!!)^{-1}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}\partial z^{2n}+o(\sigma^{2p}),$$

and the result follows.

Corollary 1.3.3.

For the 'Difference of Gaussians' (DOG) operator of Marr-Hildreth:

$$I(X)\otimes[G(X,\sigma_e)-G(X,\sigma_i)]-\sigma_e{}^2(1-r^2)\nabla^2 I(X)/2$$
$$=\sigma_e{}^4(1-r^4)\nabla^4 I(X)/8+o(\sigma_e{}^4),\ \sigma_e\to 0.$$

Note that the approximant in the thorem is nothing but σ^{-2}-times the error in the Gaussian filtering which is separable and can be obtained using three 1-D convolutions. Moreover, the asymptotic error is ½-times the corresponding error in the approximation of the Laplacian by the Marr-Hildreth operator. This formulation for the Laplacian gives good results if $I(X)$ is a reasonably noise free float reconstruction. However, if $I(X)$ is a noisy or a low order digitized version, a further filtering with a $G(X,\sigma_0)$ could correct the situation. Since

$$I(X)\otimes G(X,\sigma_i)\otimes G(X,\sigma_0)=I(X)\otimes G(X,\sigma_e),$$

where $\sigma_e^2 = \sigma_i^2 + \sigma_0^2$, (K. Yoshida [1971, p. 235]) one is lead to another formulation for the Difference of Gaussians (DOG) operator:

$$[G(X,\sigma_e) - G(X,\sigma_i)] \otimes I(X) = G(X,\sigma_i) \otimes [G(X,\sigma_0) \otimes I(X) - I(X)].$$

1.4. COMPUTING GRADIENTS OF ARBITRARY ORDER

A class $I_\sigma(X)$ of functions is called an approximation of $I(X)$. if in some sense (e.g., pointwise, uniform, L^p, etc.) $I_\sigma(X) \to I(X), \sigma \to 0$. If for all non-negative integers r, s, $t \geq 0$, $\partial^{r+s+t} I_\sigma(X)/\partial x^r \partial y^s \partial z^t \to \partial^{r+s+t} I(X)/\partial x^r \partial y^s \partial z^t$, as $\sigma \to 0$, $I_\sigma(X)$ is said to possess simultaneous approximation property of an arbitrary order. We show in the following that a convolution with the gradients (a linear combination of partial derivatives of certain orders) of the Gaussian filter function provides pointwise approximation to the corresponding gradients of $I(X)$ of a similar order of accuracy as obtained in the previous approximations. The following result is crucially needed in proving that $I_\sigma(X) = I(X) \otimes G(X,\sigma)$ does possess the pointwise simultaneous approximation property of an arbitrary order:

Lemma 1.4.1.

For $k \geq 1$,

$$\partial^k G(x,\sigma)/\partial x^k = \sum_{i,j} c_{ij}^{[k]} (x^j/\sigma^i) G(x,\sigma),$$

where $c_{ij}^{[k]}$ are constants and the sum runs over $i, j \geq 0$, such that $i-j \leq k$.

Proof: The result is true for $k = 1$, for, $\partial G(x,\sigma)/\partial x = -(x/\sigma^2) G(x,\sigma)$. Hence, assuming the result for k,

$$\partial^{k+1} G(x,\sigma)/\partial x^{k+1} = \sum_{i,j} c_{ij}^{[k]} (jx^{j-1}/\sigma^i - x^{j+1}/\sigma^{i+2}) G(x,\sigma),$$

which is of the given type.

Theorem 1.4.2.

Let $I(X)$ be a bounded continuous function, q-differentiable at a point X If $Q(u,v,w)$ is any polynomial of total degree q in u, v, w, then with $Q(D)$ denoting the differential operator $Q(\partial/\partial x, \partial/\partial y, \partial/\partial z)$,

$$\lim_{\sigma\to 0} I(X)\otimes[Q(D)G(X,\sigma)] = Q(D)I(X),$$

the relation holding uniformly in $X \in S$ if $I(X)$ is uniformly q-differentiable in S. Moreover, if $I(X)$ is $2p+q$-differentiable at X, then

$$Q(D)[I(X)\otimes G(X,\sigma) - I(X)] = \sum_k \sigma^{2k} \sum_{\lambda,\mu,\nu} ((2\lambda)!!(2\mu)!!(2\nu)!!)^{-1} \times \left(\partial^{2k}[Q(D)I(X)]/\partial x^{2\lambda}\partial y^{2\mu}\partial z^{2\nu}\right) + o(\sigma^{2p}),$$

where the sums run over $2 \le k \le p$, $\lambda,\mu,\nu \ge 0$, with $\lambda+\mu+\nu = k$, and the relation holds uniformly in $X \in S$ if $I(X)$ is uniformly $2p+q$-differentiable in S.

Proof: It is enough to consider the case $Q(D) = \partial^{l+m+n}/\partial x^l \partial y^m \partial z^n$, where l, m, $n \ge 0$, and $l+m+n \le q$. Since q-differentiability implies $l+m+n$-differentiability of $I(X)$ at X,

$$I(X)\otimes Q(D)G(X,\sigma) = \iiint I(X-H)Q(D)G(H,\sigma)dH$$

$$= \iiint [I(X) + \sum_{1\le k\le l+m+n} (k!)^{-1}(-H\cdot\nabla)^k I(X) + o(H^{l+m+n})]\, Q(D)G(H,\sigma)dH.$$

Using integration by parts,

$$\int_R x^k \partial^{l+m+n} G(X,\sigma)/\partial x^l \partial y^m \partial z^n dx = -\int_R k x^{k-1} \partial^{l+m+n-1} G(X,\sigma)/\partial x^{l-1}\partial y^m \partial z^n dx,$$

as the integrated part equals zero. It follows that

$$\int_{-\infty}^{\infty} x^k \frac{\partial^{l+m+n} G(X,\sigma)}{\partial x^l \partial y^m \partial z^n} dx = \begin{cases} 0, \text{if } k < l, \\ (-1)^l l! \int_{-\infty}^{\infty} \frac{\partial^{m+n} G(X,\sigma)}{\partial y^m \partial z^n} dx = l! G^{(m)}(y,\sigma) G^{(n)}(z,\sigma), \\ \qquad\qquad \text{if } k = l, \\ (-1)^l \frac{k!}{(k-l)!} \left(\int_{-\infty}^{\infty} x^{k-l} G(x,\sigma) dx\right) G^{(m)}(y,\sigma) G^{(n)}(z,\sigma), \\ \qquad\qquad \text{if } k > l. \end{cases}$$

Further, if $k > l$,

$$\int_{-\infty}^{\infty} x^k \frac{\partial^{l+m+n} G(X,\sigma)}{\partial x^l \partial y^m \partial z^n} dx = \begin{cases} 0, \text{if } (k-l) \text{ is odd}, \\ (-1)^l \dfrac{k!}{(k-l)!!} \sigma^{k-l} G^{(m)}(y,\sigma) G^{(n)}(z,\sigma), \text{if } (k-l) \text{ is even}. \end{cases}$$

Similar relations are valid with x, y, z cyclically interchanged. It follows that

$$I(X) \otimes Q(D) G(X,\sigma) = ((l+m+n)!)^{-1} l!m!n![(l+m+n)!/(l!m!n!)] Q(D) I(X)$$

$$+ \int_{\cdots}^{\cdots} [o(|H|^{l+m+n})] Q(D) G(H,\sigma) dH$$

$$= Q(D) I(X) + \int_{\cdots}^{\cdots} [o(H^{l+m+n})] Q(D) G(H,\sigma) dH.$$

Now, given an arbitrary $\varepsilon > 0$, there exists a $\delta > 0$, such that

$$o(H^{l+m+n}) < \left(\varepsilon + H^2/\delta^2\right) H^{l+m+n}, H \in \mathrm{R}^3,$$

Hence, taking $H = (\xi, \eta, \varsigma)'$, we have from the Lemma that

$$\int_{\cdots}^{\cdots} [o(H^{l+m+n})] Q(D) G(H,\sigma) dH$$

$$\le \int_{\cdots}^{\cdots} [(\varepsilon + H^2/\delta^2) H^{l+m+n}] \sum_{r,s \ge 0, r-s \le l} c_{rs}^{[l]} (\xi^s/\sigma^r) \sum_{t,u \ge 0, t-u \le m} c_{tu}^{[m]} (\eta^u/\sigma^t) \cdot$$

$$\times \sum_{v,w \ge 0, v-w \le n} c_{vw}^{[n]} (\zeta^w/\sigma^v) \, G(H,\sigma) dH$$

$$\le \varepsilon O\left(\sum_{r,s \ge 0, r-s \le l,\ t,u \ge 0, t-u \le m,\ v,w \ge 0, v-w \le n} \sigma^{l+m+n-(r-s)-(t-u)-(v-w)}\right)$$

$$+ O\left(\sum_{r,s \ge 0, r-s \le l,\ t,u \ge 0, t-u \le m,\ v,w \ge 0, v-w \le n} \delta^{-2} \sigma^{2+l+m+n-(r-s)-(t-u)-(v-w)}\right)$$

$$\le \varepsilon O(1) + O(\sigma^2/\delta^2).$$

Due to the arbitrariness of $\varepsilon > 0$, it follows that the contribution of the o-term goes to zero as $\sigma \to 0$, and the required limit relation

$$\lim_{\sigma \to 0} I(X) \otimes [Q(D) G(X,\sigma)] = Q(D) I(X)$$

follows.

If $I(X)$ is uniformly q-differentiable in S, the $[\varepsilon O(1) + O(\sigma^2/\delta^2)]$- term in the above is independent of $X \in S$, and the uniformity of the limit relation follows.

If $I(X)$ is $2p+q$-differentiable at X, we have

$$Q(D)[I(X) \otimes G(X,\sigma) - I(X)]$$

$$= \int_{\cdots}^{\cdots} [\sum_{0 \le k \le q+2p} (\mathrm{k}!)^{-1} (-H \cdot \nabla)^k I(X) + o(|H|^{q+2p})] Q(D) G(H,\sigma) dH$$

$$= Q(D)I(X) + \int_{\cdots}^{\cdots}[\sum_{l+m+n+1\le k\le q+2p}(\mathrm{k!})^{-1}(-H\cdot\nabla)^k I(X) + \mathrm{o}(|H|^{q+2p})]\,Q(D)G(H,\sigma)dH.$$

We note that for $k > l+m+n$, and $r+s+t=k$,

$$\int_{\cdots}^{\cdots} x^r y^s z^t [\partial^{\lambda+\mu+\nu} G(X,\sigma)/\partial x^\lambda \partial y^\mu \partial z^\nu]dX$$

$$= (-1)^{r+s+t}[r!s!t!/\{(r-\lambda)!!(s-\mu)!!(t-\nu)!!\}]\sigma^{r+s+t-(\lambda+\mu+\nu)},$$

provided $r-\lambda, s-\mu, t-\nu$ are all even and non-negative integers, and a vacuous semi-factorial is to be interpreted as 1, which arises whenever any of $r-\lambda, s-\mu, t-\nu$ is zero.

Also, if any of $r-\lambda, s-\mu, t-\nu$ is odd or negative

$$\int_{\cdots}^{\cdots} x^r y^s z^t [\partial^{\lambda+\mu+\nu} G(X,\sigma)/\partial x^\lambda \partial y^\mu \partial z^\nu]dX = 0.$$

Hence,

$$\int_{\cdots}^{\cdots} \sum_{l+m+n+1\le k\le q+2p} (k!)^{-1}[(-H\cdot\nabla)^k I(X)]\,Q(D)G(H,\sigma)dH$$

$$= \sum_{l+m+n+1\le k\le q+2p}(k!)^{-1}\sum_{2(\lambda+\mu+\nu)=k-(l+m+n)}\{[(l+2\lambda)!(m+2\mu)!(n+2\nu)!]$$

$$\times[(2\lambda)!!(2\mu)!!(2\nu)!!]^{-1}\}\sigma^{k-(l+m+n)}$$

$$\times[k!/\{(l+2\lambda)!(m+2\mu)!(n+2\nu)!\}]\partial^{2(\lambda+\mu+\nu)}Q(D)I(X)/\partial x^{2\lambda}\partial y^{2\mu}\partial z^{2\nu}$$

$$= \sum_{1\le j\le q+2p-(l+m+n)}\sigma^j\sum_{2(\lambda+\mu+\nu)=j}[(2\lambda)!!(2\mu)!!(2\nu)!!]^{-1}\partial^{2j}Q(D)I(X)/\partial x^{2\lambda}\partial y^{2\mu}\partial z^{2\nu}$$

$$= \sum_{1\le j\le p}\sigma^{2j}\sum_{(\lambda+\mu+\nu)=j}[(2\lambda)!!(2\mu)!!(2\nu)!!]^{-1}\partial^{2j}Q(D)I(X)/\partial x^{2\lambda}\partial y^{2\mu}\partial z^{2\nu} + \mathrm{o}(\sigma^{2p}).$$

That the contribution of the $\mathrm{o}(|H|^{q+2p})$-term is $\mathrm{o}(\sigma^{2p})$ is clear from:

$$\int_{\cdots}^{\cdots}\left(\mathrm{o}(|H|^{l+m+n})\right)Q(D)G(H,\sigma)dH$$

$$\le \int_{\cdots}^{\cdots}[(\varepsilon + |H|^2/\delta^2)|H|^{q+2p}]\sum_{r,s\ge 0,r-s\le l} c_{rs}^{[l]}(\xi^s/\sigma^r)\sum_{t,u\ge 0,t-u\le m} c_{tu}^{[m]}(\eta^u/\sigma^t)$$

$$\times\sum_{v,w\ge 0,v-w\le n} c_{vw}^{[n]}(\zeta^w/\sigma^v)\,G(H,\sigma)dH$$

$$\le \varepsilon\mathrm{O}\Big(\sum_{r,s\ge 0,r-s\le l,\ t,u\ge 0,t-u\le m,\ v,w\ge 0,v-w\le n}\sigma^{q+2p-(r-s)-(t-u)-(v-w)}\Big)$$

$$+\mathrm{O}\Big(\sum_{r,s\ge 0,r-s\le l,\ t,u\ge 0,t-u\le m,\ v,w\ge 0,v-w\le n}\delta^{-2}\sigma^{2+q+2p-(r-s)-(t-u)-(v-w)}\Big)$$

$$\le \varepsilon\mathrm{O}(\sigma^{2p}) + \mathrm{O}(\sigma^{2p+2}/\delta^2),$$

since $2p+q-(l+m+n)+l-(r-s)+m-(t-u)+n-(v-w) \geq 2p+q-(l+m+n) \geq 0$. If $I(X)$ is uniformly $2p+q$-differentiable in S, the uniformity assertion for $X \in S$ follows as in this case δ can be chosen to be independent of $X \in S$, in which case the $[\varepsilon O(\sigma^{2p}) + O(\sigma^{2p+2}/\delta^2)]$ - term does not depend on $X \in S$. This completes the proof.

1.5. ERROR ANALYSIS FOR SMOOTH GRADIENT FUNCTIONS

Quite an insight into the approximations of gradient functions could be obtained by assuming that the images under consideration possess sufficiently smooth gradient functions. In fact the images obtained in MRI using FFT on finite raw data could be considered infinitely smooth - they being trigonometric polynomials in the variables x, y and z. In this section we obtain some inverse results on their degree of approximation, which for a bounded continuous gradient $Q(D) = Q(\partial/\partial x, \partial/\partial y, \partial/\partial z) I(X)$ to a large extent, characterizes the smoothness of the gradient. Some definitions needed follow: Let C denote the Banach space of all bounded continuous functions on R^3 normed by

$$I = \sup_{X \in R^3} \{I(X)\},$$

and let C^2 denote the Banach space of bounded 2-differentiable functions $I(X)$ on R^3, having all partial derivatives upto order 2 bounded on R^3. The space C^2 could be normed by $I_{C,2} = I + I_2$, where we define

$$|I_2 = \max\{\|I_{xx}\|, I_{yy}, \|I_{zz}|, (4/\pi) I_{yz}, (4/\pi)\|I_{zx}\|, (4/\pi) I_{xy}\}$$

is a semi-norm on C^2. Here the constants $(4/\pi)$ have been chosen so as to simplify some expressions later.

Associated with the spaces C and C^2, we define the Peetre's K-functional by:

$$K(I(X);t) = \inf_{g \in C^2} \{ I - g + t^2 g(X)_2 \} I(X) \in C.$$

It may be noted that for $f \in C$,

$$I(X) \otimes G(X,\sigma) - I(X) \leq \int\int\int [I(X - \sigma H) - I(X)] G(H,1) dH .$$

Hence, if $I(X)$ is continuous on R^3, the integrand being dominated by $2\,I\,G(H,1)$ and tending to zero pointwise as $\sigma \to 0$, it follows by Lebesgue's dominated convergence theorem that $I(X) \otimes G(X,\sigma) \to I(X), \sigma \to 0$. Further, this pointwise convergence is easily seen to be uniform on each compact subset of R^3 (H.S. Shapiro [1969, p. 14]). If, however, $I(X)$ is uniformly continuous on R^3.

$$\omega(\delta) = \sup\{|I(X+H) - I(X)| : |H| \le \delta\} \to 0, \delta \to 0,$$

and then the integrand being dominated by $\omega(\delta H)G(H,1)$, it follows that the convergence $I(X) \otimes G(X,\sigma) \to I(X), \sigma \to 0$, is uniform on R^3.

The degree of approximation of C^2-functions by a Gaussian convolution is estimated in the next result, which will also be used in a subsequent result relating the degree of approximation with the smoothness of the function being approximated.

Lemma 1.5.1.

$$I(X) \otimes G(X,\sigma) - I(X) \le 3\sigma^2 \|I\|_2, I \in C^2.$$

Proof: If $I(X) \in C^2$, by the mean value theorem of vector calculus,

$$I(X+H) - I(X) - H \cdot \nabla I(X) \le (1/2)[h_1^2 I_{xx} + h_2^2 I_{yy} + h_3^2 I_{zz} + 2h_2h_3 I_{yz} + 2h_3h_1 I_{zx} + 2h_1h_2 I_{xy}](X + \theta H),$$

where $0 < \theta < 1$, and all the mixed partial derivatives have been evaluated at the indicated point $X+\theta H$. It follows that

$$\begin{aligned} I(X) \otimes G(X,\sigma) - I(X) &\le (1/2) \iiint [h_1^2 I_{xx} + h_2^2 I_{yy} + h_3^2 I_{zz} \\ &\quad + 2h_2 h_3 I_{yz} + 2h_3 h_1 |I_{zx} + 2h_1 h_2 I_{xy}] G(H,\sigma) dH \\ &\le (I_2/2) \iiint [h_1^2 + h_2^2 + h_3^2 + (\pi/2)(h_2 h_3 + h_3 h_1 + h_1 h_2)] G(H,\sigma) dH \\ &\le 3 I_2 \sigma^2, \end{aligned}$$

where we have used the following evaluations (R.K.S. Rathore [1973, p. 65]):

$$(2\pi)^{-1/2} \sigma^{-1} \int t \exp\{-t^2/(2\sigma^2)\} dt = (2/\pi)^{1/2} \sigma,$$

Hence, if $I(X)$ is continuous on R^3, the integrand being dominated by $2\,I\,G(H,1)$ and tending to zero pointwise as $\sigma \to 0$, it follows by Lebesgue's dominated convergence theorem that $I(X) \otimes G(X,\sigma) \to I(X), \sigma \to 0$. Further, this pointwise convergence is easily seen to be uniform on each compact subset of R^3 (H.S. Shapiro [1969, p. 14]). If, however, $I(X)$ is uniformly continuous on R^3.

$$\omega(\delta) = \sup\{|I(X+H) - I(X)| : H \le \delta\} \to 0, \delta \to 0,$$

and then the integrand being dominated by $\omega(\delta H)G(H,1)$, it follows that the convergence $I(X) \otimes G(X,\sigma) \to I(X), \sigma \to 0$, is uniform on R^3.

The degree of approximation of C^2-functions by a Gaussian convolution is estimated in the next result, which will also be used in a subsequent result relating the degree of approximation with the smoothness of the function being approximated.

Lemma 1.5.1.

$|I(X) \otimes G(X,\sigma) - I(X)| \le 3\sigma^2 \|I\|_2, I \in C^2.$

Proof: If $I(X) \in C^2$, by the mean value theorem of vector calculus,

$$I(X+H) - I(X) - H \cdot \nabla I(X) \le (1/2)[h_1^2 I_{xx} + h_2^2 I_{yy} + h_3^2 I_{zz} + 2h_2h_3I_{yz} + 2h_3h_1I_{zx} + 2h_1h_2I_{xy}](X + \theta H),$$

where $0 < \theta < 1$, and all the mixed partial derivatives have been evaluated at the indicated point $X+\theta H$. It follows that

$$\begin{aligned} I(X) \otimes G(X,\sigma) - I(X) &\le (1/2)\int\cdots\int [h_1^2 I_{xx} + h_2^2 I_{yy} + h_3^2 I_{zz} \\ &\quad + 2h_2 h_3 I_{yz} + 2h_3 h_1 I_{zx} + 2h_1 h_2 I_{xy}]G(H,\sigma)dH \\ &\le (\|I\|_2/2)\int\cdots\int [h_1^2 + h_2^2 + h_3^2 + (\pi/2)(h_2 h_3 + |h_3 h_1 + h_1 h_2|)]G(H,\sigma)dH \\ &\le 3\|I\|_2\sigma^2, \end{aligned}$$

where we have used the following evaluations (R.K.S. Rathore [1973, p. 65]):

$$(2\pi)^{-1/2}\sigma^{-1}\int_0^\infty t \exp\{-t^2/(2\sigma^2)\}dt = (2/\pi)^{1/2}\sigma,$$

and

$$(2\pi)^{-1/2}\sigma^{-1}\int_{\bullet}^{\bullet} t^2 \exp\{-t^2/(2\sigma^2)\}dt = \sigma^2.$$

This completes the proof of the Lemma.

Lemma 1.5.2.

$$\|I(X)\otimes G(X,\sigma)\|_2 \le 2\sigma^{-2}\|I\|, I\in C.$$

Proof: Using the rules for differentiating a convolution,

$$\begin{aligned}[I(X)\otimes G(X,\sigma)]_{xx} &= \int_{\bullet\bullet\bullet}^{\bullet\bullet\bullet}[I(X-H)G_{xx}(H,\sigma)]dH \\ &= \int_{\bullet\bullet\bullet}^{\bullet\bullet\bullet}[I(X-H)\{-1/\sigma^2+(h_1/\sigma^2)^2\}G(H,\sigma)]dH \\ &\le 2\|I\|\sigma^{-2}.\end{aligned}$$

Also,

$$\begin{aligned}[I(X)\otimes G(X,\sigma)]_{xy} &= \int_{\bullet\bullet\bullet}^{\bullet\bullet\bullet} I(X-H)G_{xy}(H,\sigma)dH \\ &= \int_{\bullet\bullet\bullet}^{\bullet\bullet\bullet}[I(X-H)\{xy/\sigma^4\}G(H,\sigma)]dH \\ &\le (2/\pi)\,I\,\sigma^{-2}.\end{aligned}$$

Using symmetry, also we have:

$$[I(X)\otimes G(X,\sigma)]_{yy}\ ,\ [I(X)\otimes G(X,\sigma)]_{zz} \le 2\,I\,\sigma^{-2},\ \text{and}$$

$$[I(X)\otimes G(X,\sigma)]_{yz}\ ,\ [I(X)\otimes G(X,\sigma)]_{zx} \le (2/\pi)\,I\,\sigma^{-2}.$$

Hence,

$$I(X)\otimes G(X,\sigma)_{\ 2} \le \max\{2\,I\,\sigma^{-2},(8/\pi^2)\,I\,\sigma^{-2}\} = 2\sigma^{-2}\,I\ ,$$

completing the proof of the Lemma.

An estimate of error in approximation by a Gaussian convolution in terms of the Peetre's *K*-functional is given in the next result:

Lemma 1.5.3.

$$I(X)\otimes G(X,\sigma) - I(X) \le 3K(I;\sigma), I\in C.$$

Proof: Let $I(X)\in C$. Then, for $g\in C^2$, we have:

$$I(X)\otimes G(X,\sigma)-I(X) \le I(X)-g(X) + g(X)-g(X)\otimes G(X,\sigma) + [g(X)-I(X)]\otimes G(X,\sigma)$$
$$\le 2\|I(X)-g(X)\|+3\sigma^2\|g\|_2.$$

Taking infimum over $g\in C^2$, we have $I(X)\otimes G(X,\sigma)-I(X) \le 3K(I;\sigma)$, completing the proof.

Theorem 1.5.4.

Let $I(X)\in C$, and $\psi\in S$. Then:

(i) $I(X)\otimes G(X,\sigma)-I(X)\| = O(\psi(\sigma^2)), \sigma\to 0$, iff $K(I;t)=O(\psi(t^2)), t\to 0$;

(ii) $I(X)\otimes G(X,\sigma)-I(X) = o(\psi(\sigma^2)), \sigma\to 0$, iff $K(I;t)=o(\psi(t^2)), t\to 0$; and,

(iii) $I(X)\otimes G(X,\sigma)-I(X) \sim K(I;\sigma)$, if $K(I;t^{1/2})\in S$.

Proof: That there holds

$I(X)\otimes G(X,\sigma)-I(X) = O(\psi(\sigma^2)), \sigma\to 0$, if $K(I;t)=O(\psi(t^2)), t\to 0$, and that

$I(X)\otimes G(X,\sigma)-I(X) = o(\psi(\sigma^2)), \sigma\to 0$, if $K(I;t)=o(\psi(t^2)), t\to 0$, follows from the previous Lemma.

Conversely, if $I(X)\otimes G(X,\sigma)-I(X) = O(\psi(\sigma^2)), \sigma\to 0$, then

$$K(I;t)\le K([I(X)-I(X)\otimes G(X,\sigma)];t)+K([I(X)-g(X)]\otimes G(X,\sigma);t) + K(g(X)\otimes G(X,\sigma);t)$$
$$\le O(\psi(\sigma^2))+2(t/\sigma)^2\ I-g\ +t^2\ g\ _2.$$

Taking infimum over $g\in C^2$, $K(I;t)\le O(\psi(\sigma^2))+2(t/\sigma)^2K(I;\sigma)$, and the rest of the results follow from Lemma 1.2.2.

Note that part (iii) of the Theorem justifies the Peetre's K-functional error bound given in the previous Lemma. The symbol '~' in above, and in the sequel, refers to the two quantities being of precisely the same order: there exist positive constants A and B such that

$I(X)\otimes G(X,\sigma)-I(X) \le AK(f;\sigma)$, and, $I(X)\otimes G(X,\sigma)-I(X) \le BK(f;\sigma)$,

for all σ sufficiently small. We also note that the familiar Bernstein-Lipschitz orders $\psi(t)=t^{\alpha/2}, 0<\alpha<2$, belong to the class S. This immediately leads to the following:

Corollary 1.5.5.

Let $I \in C$, and $0 < \alpha < 2$. Then:

(i) $\left\| I(X) \otimes G(X,\sigma) - I(X) \right\| = O(\sigma^{\alpha}), \sigma \to 0$, iff $K(I;t) = O(t^{\alpha}), t \to 0$;

(ii) $\left\| I(X) \otimes G(X,\sigma) - I(X) \right\| = o(\sigma^{\alpha}), \sigma \to 0$, iff $K(I;t) = o(t^{\alpha}), t \to 0$; and,

(iii) $\left| I(X) \otimes G(X,\sigma) - I(X) \right. \sim K(I;\sigma)$, if $K(I;t) \sim t^{\alpha}$.

Let $Q(u,v,w)$, as before, be a polynomial in u, v, w and let $Q(D) \equiv Q(\partial/\partial x, \partial/\partial y, \partial/\partial z)$. Then, using $I(X) \otimes [Q(D)G(X,\sigma)] = Q(D)[I(X) \otimes G(X,\sigma)]$, the previous Theorem and the Corollary imply the following characterizations in the smooth gradient approximation by convolutions of an image by the corresponding gradients of the Gaussian function:

Theorem 1.5.6.

Let $Q(D)I(X) \in C$ and $\psi \in S$. Then:

(i) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = O(\psi(\sigma^2)), \sigma \to 0$,

iff $K(Q(D)I(X);t) = O(\psi(t^2)), t \to 0$;

(ii) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = o(\psi(\sigma^2)), \sigma \to 0$,

iff $K(Q(D)I(X);t) = o(\psi(t^2)), t \to 0$; and,

(iii) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) \sim K(I;\sigma)$, if $K(Q(D)I(X);t^{1/2}) \in S$.

Corollary 1.5.7.

Let $Q(D)I(X) \in C$, and $0 < \alpha < 2$. Then:

(i) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) \left. \right| = O(\sigma^{\alpha}), \sigma \to 0$,

iff $K(Q(D)I(X);t) = O(t^{\alpha}), t \to 0$;

(ii) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) \left\| \right. = o(\sigma^{\alpha}), \sigma \to 0$,

iff $K(Q(D)I(X);t) = o(t^{\alpha}), t \to 0$; and,

(iv) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) \sim K(I;\sigma)$, if $K(Q(D)I(X);t) \sim t^{\alpha}$.

1.6. SATURATION OF CONVOLUTIONS WITH GAUSSIAN FILTERS

Finally, we complete the saturation theory for the convolution approximation with the Gaussian filter. Let the radial average $I_\rho(X)$ of a function $I(X) \in C$ be defined as:

$$I_\rho(X) = (4\pi\rho^2)^{-1} \iint_{A(X,\rho)} I(T)dA(T), \rho > 0,$$

where $A(X,\rho)$ denotes the surface of the sphere about the point X of radius ρ, T is a general point and $dA(T)$ stands for the area element on the surface about the point T. We call $I(X)$ harmonic if $I_\rho(X) = I(X)$, $X \in \mathrm{R}^3$, $\rho > 0$.

Theorem 1.6.1.

Let $I(X) \in C$. Then:

(i) $I(X) \otimes G(X,\sigma) - I(X) = \mathrm{O}(\sigma^2), \sigma \to 0$, iff $I_r(X) - I(X) = \mathrm{O}(r^2), r \to 0$;

and,

(ii) $I(X) \otimes G(X,\sigma) - I(X) = \mathrm{o}(\sigma^2), \sigma \to 0$, iff $I(X)$ is harmonic.

As an immediate consequence of this Theorem, we have the following:

Corollary 1.6.2.

Let $Q(D)I(X) \in C$. Then:

(i) $|I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = \mathrm{O}(\sigma^2), \sigma \to 0,$

iff $Q(D)I_r(X) - Q(D)I(X) = \mathrm{O}(r^2), r \to 0$; and,

(ii) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = \mathrm{o}(\sigma^2), \sigma \to 0,$

iff $Q(D)I(X)$ is harmonic.

Thus the convolutions with $G(X,\sigma)$ and its gradients are saturated with the order σ^2, the saturation classes being functions and gradients uniformly approximable by their radial averages with order r^2, and the trivial classes being the

classes of harmonic functions and harmonic gradients. A proof of the Theorem needs the following:

Lemma 1.6.3.

If $f(X)$ possesses continuous second order partial derivatives on R^3, and, if $\nabla^2 f \in C$, then:

$$f_\rho - f \le \rho^2 \left| \nabla^2 f \right| / 6, \rho > 0,$$

where equality holds if $f(X) \equiv |X|^2$.

Proof: We have,

$$\begin{aligned} f_\rho(X) - f(X) &= (4\pi\rho^2)^{-1} \iint_{A(X,\rho)} [f(T) - f(X)] dA(T) \\ &= (4\pi)^{-1} \iint_{A(0,1)} [f(X - \rho T) - f(X)] dA(T) \\ &= (4\pi)^{-1} \int_{[0,\rho]} \left(r^{-2} \iint_{A(X,r)} [\partial f(T)/\partial n] dA(T) \right) dr, \end{aligned}$$

By the Gauss' divergence theorem,

$$f_\rho(X) - f(X) = (4\pi)^{-1} \int_{[0,\rho]} \left(r^{-2} \iiint_{S(X,r)} [\nabla^2 f(T)] dV(T) \right) dr,$$

where $S(X,r)$ denotes the sphere in R^3 with center at X and radius r, and $dV(T)$ stands for the volume element. Hence

$$\begin{aligned} f_\rho(X) - f(X) &\le (4\pi)^{-1} \int_{[0,\rho]} \left(r^{-2} \iiint_{S(X,r)} dV(T) \right) dr \; \nabla^2 f(T) \\ &= \rho^2 \nabla^2 f(X) / 6, \end{aligned}$$

completing the proof of the Lemma.

Proof of the Theorem: Let us first assume that $I_r(X) - I(X) = O(r^2), r \to 0$. Then, as $I(X) \in C$, $I_r(X) - I(X) \le Mr^2, r > 0$, for some $M > 0$. Using Fubini's theorem,

$$\begin{aligned} I(X) \otimes G(X,\sigma) - I(X) &= \iiint [I(X - T) - I(X)] G(T,\sigma) dT \\ &= \int_{[0,\infty)} \left(\iint_{A(0,\rho)} [I(X - T) - I(X)] G(T,\sigma) dA(T) \right) d\rho \\ &= \int_{[0,\infty)} 4\pi\rho^2 [I_\rho(X) - I(X)] (2\pi\sigma^2)^{-3/2} \exp\{-\rho^2/2\sigma^2\} d\rho. \end{aligned}$$

Hence,

$$I(X)\otimes G(X,\sigma)-I(X)\le M\int_{[0,\infty)}2(\rho^4/\sigma^2)(2\pi\sigma^2)^{-3/2}\exp\{-\rho^2/2\sigma^2\}d\rho$$
$$=3M\sigma^2=\mathrm{O}(\sigma^2).$$

Conversely, if $|I(X)\otimes G(X,\sigma)-I(X)|=\mathrm{O}(\sigma^2)$, writing:

$$I^{[\tau]}(X)=I(X)\otimes G(X,\tau),$$

we have, $|I^{[\tau]}(X)\otimes G(X,\sigma)-I^{[\tau]}(X)|=\mathrm{O}(\sigma^2)$. Hence, for some $B>0$,

$$I^{[\tau]}(X)\otimes G(X,\sigma)-I^{[\tau]}(X)\le B\sigma^2,$$

for all σ sufficiently small. By the asymptotic formula for the Gaussian filter

$$I^{[\tau]}(X)\otimes G(X,\sigma)-I^{[\tau]}(X)=(\sigma^2/2)\nabla^2 I^{[\tau]}(X)+\mathrm{o}(\sigma^2).$$

Hence, $\nabla^2 I^{[\tau]}(X)\le 2B$, and by the previous Lemma,

$$(I^{[\tau]})_\rho(X)-I^{[\tau]}(X)\le\rho^2\ \nabla^2 I^{[\tau]}(X)\big/6\le B\rho^2/3.$$

Taking limit as $\tau\to 0$, we get: $I_\rho(X)-I(X)\le B\rho^2/3=\mathrm{O}(\rho^2)$, which proves part (i). For (ii), if $I(X)$ is harmonic, $0=|I(X)\otimes G(X,\sigma)-I(X)|=\mathrm{o}(\sigma^2)$.

Conversely, from the latter, $|I^{[\tau]}(X)\otimes G(X,\sigma)-I^{[\tau]}(X)|=\mathrm{o}(\sigma^2)$, uniformly in τ. Hence, given any $\varepsilon>0$, there is a $\sigma_0>0$ such that

$$I^{[\tau]}(X)\otimes G(X,\sigma)-I^{[\tau]}(X)\le\varepsilon\sigma^2,0<\sigma<\sigma_0.$$

It follows that: $\nabla^2 I^{[\tau]}(X)\le 2\varepsilon$, and hence by the previous Lemma that

$$(I^{[\tau]})_\rho(X)-I^{[\tau]}(X)\le\varepsilon\rho^2/3.$$

Letting $\tau\to 0$, $I_\rho(X)-I(X)\le\varepsilon\rho^2/3,\rho>0$. As ε is arbitrary $I_\rho(X)\equiv I(X)$, completing the proof.

1.7. 2-D CONVOLUTIONS WITH THE GAUSSIAN FUNCTION

Results similar to the 3-D case remain valid for the 2-D case of convolutions with the Gaussian function. In this section we consider such results. In view of the

detailed proofs in the 3-D case, here we mostly outline the proofs to point out the differences, if any. In the 2-D case, $X \equiv (x, y)'$ denotes a point in R^2,

$$\nabla \equiv [i(\partial/\partial x + j(\partial/\partial y)], \nabla^2 f(X) = (\partial^2 f/\partial x^2 + \partial^2 f/\partial y^2), \text{ and}$$

$$\nabla^4 f(X) = (\partial^2 f/\partial x^2 + \partial^2 f/\partial y^2)^2.$$

A bounded continuous function $f(X)$ for which:

$$f(X+H) = f(X) + \sum_{1\le k\le m} (k!)^{-1}(H\cdot\nabla)^k f(X) + o(|H^m|), |H| \to 0,$$

is called m-differentiable at a point $X \in R^2$; $f(X)$ is called uniformly m-differentiable on a set $S \subset R^2$, if given an arbitrary $\varepsilon > 0$, there exists a $\delta > 0$, independent of X, such that whenever $H < \delta$,

$$f(X+H) - f(X) + \sum_{1\le k\le m} (k!)^{-1}(H\cdot\nabla)^k f(X) < \varepsilon H^m, X \in S.$$

For the 2-D Gaussian filter

$$G(X,\sigma) = (2\pi\sigma^2)^{-1}\exp[-X^2/(2\sigma^2)] = (2\pi\sigma^2)^{-1}\exp[-(x^2+y^2)/(2\sigma^2)],$$

there holds the following asymptotic error estimate:

Theorem 1.7.1.

If $I(X)$ is 4-differentiable at a point X,

$$\lim_{\sigma\to 0} \sigma^{-2}\{I(X)\otimes\nabla^2 G(X,\sigma) - \nabla^2 I(X)\} = \nabla^4 I(X)/2,$$

and this relation holds uniformly in $X \in S \subset R^2$, if $I(X)$ is uniformly 4-differentiable in S. If $I(X)$ is $2p$-differentiable at a point X,

$$I(X)\otimes\nabla^2 G(X,\sigma) - \nabla^2 I(X) = \sum_k (2k)\sigma^{2(k-1)} \sum_{l,m,n} ((2l)!!(2m)!!)^{-1} \times (\partial^{2k} I(X)/\partial x^{2l}\partial y^{2m}) + o(\sigma^{2(p-1)}),$$

where the sums run over $2 \le k \le p$, l, m, $n \ge 0$, with $l+m+n = k$. Moreover, the relation holds uniformly in $X \in S$ if $I(X)$ is uniformly $2p$-differentiable in S.

Proof: With the double integral $\iint [.]dH$ representing an integral over R^2, we have

$$I(X)\otimes\nabla^2 G(X,\sigma) = I(X)\otimes[(2\pi)^{-1}\sigma^{-4}\exp\{-|X^2/(2\sigma^2)\}(-2 + X^2/\sigma^2)]$$

$$= \int\int [I(X) + \sum_{1\le k\le 2p} (k!)^{-1}(-H\cdot\nabla)^k I(X) + o(|H|^{2p})](2\pi)^{-1}\sigma^{-4}$$

$$\times \exp\{-|H|^2/(2\sigma^2)\}(-2+|H|/\sigma^2)dH.$$

Using $(2\pi)^{-1/2}\sigma^{-1}\int t^j \exp(-t^2/2\sigma^2)dt = 0$, if j is odd and $(j\text{-}1)!!\sigma^j$, if j is even,

$$\int\int \Big(\sum_{0\le k\le 2p} (k!)^{-1}(-H\cdot\nabla)^k I(X)\Big)(2\pi)^{-1}\sigma^{-4}\exp\{-|H|^2/(2\sigma^2)\}(-2+|H|^2/\sigma^2)dH$$

$$= \sum_{0\le k\le p}((2k)!)^{-1}\sum_{l,m\ge 0, l+m=2k}[(2k)!/(l!m!)]\partial^{2k}I(X)/\partial x^l\partial y^m(2\pi)^{-1}\sigma^{-4}$$

$$\times \int\int \lambda^l\mu^m \exp\{-(\lambda^2+\mu^2)/(2\sigma^2)\}[-2+(\lambda^2+\mu^2)/\sigma^2]d\lambda d\mu$$

$$= \sum_{0\le k\le p}\sum_{l,m\ge 0, l+m=k}[1/((2l)!(2m)!)]\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}(2\pi)^{-1}\sigma^{-4}$$

$$\times \int\int \lambda^{2l}\mu^{2m} \exp\{-(\lambda^2+\mu^2)/(2\sigma^2)\}[-2+(\lambda^2+\mu^2)/\sigma^2]d\lambda d\mu$$

$$= \sum_{1\le k\le p}\sum_{l,m\ge 0, l+m=k} 2k/((2l)!!(2m)!!)\sigma^{2(l+m-1)}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}$$

$$= (\partial^2/\partial x^2 + \partial^2/\partial y^2)I(X) + \sum_{2\le k\le p}\sum_{l,m\ge 0, l+m=k} 2k/((2l)!!(2m)!!)$$

$$\times \sigma^{2(k-1)}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}.$$

For the $o(|H|^{2p})$-term, given an arbitrary $\varepsilon > 0$, there exists a $\delta > 0$, such that $o(|H|^{2p}) < (\varepsilon + |H|^2/\delta^2)|H|^{2p}$, $H \in R^3$. Using,

$$\int\int |X|^{2p+2j} G(X,\sigma)dX = O(\sigma^{2p+2j}), \sigma \to 0,$$

the net contribution of the $o(|H|^{2p})$–term being $o(\sigma^{2(p-1)})$, and the result for $2p$-differentiable $I(X)$ follows. In particular, for $p = 2$,

$$\lim_{\sigma\to 0}\sigma^{-2}[I(X)\otimes\nabla^2 G(X,\sigma) - \nabla^2 I(X)]$$

$$= 4\sum_{l,m\ge 0, l+m=2}((2l)!!(2m)!!)^{-1}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}$$

$$= 4[\{\partial^4/\partial x^4 + \partial^4/\partial y^4\}I(X)/8 + \{\partial^4/\partial y^2\partial z^2 + \partial^4/\partial z^2\partial x^2\}I(X)/4]$$

$$= (1/2)(\partial^2/\partial x^2 + \partial^2/\partial y^2)^2 I(X).$$

The uniformity of the asymptotic estimates in the cases of uniform differentiability is clear as the δ in above could be chosen to be independent of $X \in S$.

145048

Theorem 1.7.2.

If $I(X)$ is $2p$-differentiable at a point X,

$$[I(X)\otimes G(X,\sigma)-I(X)]=(\sigma^2/2)\nabla^2 I(X)$$
$$+\sum_k \sigma^{2k}\sum_{l,m}((2l)!!(2m)!!)^{-1}(\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m})+o(\sigma^{2p}),$$

the sums being over $2 \le k \le p$, with l, $m \ge 0$, and $l+m = k$. This relation holds uniformly in $X \in S \subset R^2$ if $I(X)$ is uniformly $2p$-differentiable in S. In particular, if $I(X)$ is 4-differentiable at a point X,

$$\lim_{\sigma\to 0}\sigma^{-2}\left\{2\sigma^{-2}[I(X)\otimes\nabla^2 G(X,\sigma)-I(X)]-\nabla^2 I(X)\right\}=\nabla^4 I(X)/4,$$

and it holds uniformly in $X \in S$, if $I(X)$ is uniformly 4-differentiable in S.

Proof:

$$I(X)\otimes G(X;\sigma)=I(X)\otimes\left((2\pi)^{-1}\sigma^{-2}\exp\left\{-|X|^2/(2\sigma^2)\right\}\right)$$
$$=\int\!\!\int\left\{I(X)+\sum_{1\le k\le 2p}(k!)^{-1}(-H\cdot\nabla)^k I(X)+o(|H|^{2p})\right\}(2\pi)^{-1}\sigma^{-2}$$
$$\times\exp\left\{-|H|^2/(2\sigma^2)\right\}dH.$$

As before, the $o(|H|^{2p})$-term on the right hand side contributes to $o(\sigma^{2p})$, which holds uniformly in $X \in S$, in the uniform differentiability case. For the remaining term:

$$\int\!\!\int\left(\sum_{0\le k\le 2p}(k!)^{-1}(-H\cdot\nabla)^k I(X)\right)(2\pi)^{-1}\sigma^{-2}\exp\left\{-|H|^2/(2\sigma^2)\right\}dH$$
$$=\sum_{0\le k\le p}((2k)!)^{-1}\sum_{l,m\ge 0,l+m=2k}[(2k)!/(l!m!)]\partial^{2k}I(X)/\partial x^l\partial y^m(2\pi)^{-1}\sigma^{-2}$$
$$\times\int\!\!\int\lambda^l\mu^m\exp\{-(\lambda^2+\mu^2)/(2\sigma^2)\}d\lambda d\mu$$
$$=\sum_{0\le k\le p}\sum_{l,m\ge 0,l+m=k}[1/((2l)!(2m)!)]\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}(2\pi)^{-1}\sigma^{-2}$$
$$\times\int\!\!\int\lambda^{2l}\mu^{2m}\exp\{-(\lambda^2+\mu^2)/(2\sigma^2)\}d\lambda d\mu$$
$$=\sum_{0\le k\le p}\sum_{l,m\ge 0,l+m=k}((2l)!(2m)!)^{-1}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}$$
$$\times[(2l-1)!!(2m-1)!!]\sigma^{(2l+2m)}$$
$$=\sum_{0\le k\le p}\sum_{l,m\ge 0,l+m=k}((2l)!!(2m)!!)^{-1}\sigma^{2k}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}.$$

So,

$$I(X)\otimes G(X,\sigma)=I(X)+\sigma^2(\partial^2/\partial x^2+\partial^2/\partial y^2)I(X)/2$$
$$+\sigma^4(\partial^2/\partial x^2+\partial^2/\partial y^2)^2 I(X)/8$$
$$+\sum_{3\le k\le p}\sigma^{2k}\sum_{l,m\ge 0,l+m=k}((2l)!!(2m)!!)^{-1}\partial^{2k}I(X)/\partial x^{2l}\partial y^{2m}+o(\sigma^{2p}),$$

and the result follows.

Corollary 1.7.3.

For the difference of Gaussians (DOG) in two variables:

$$I(X)\otimes[G(X,\sigma_e)-G(X,\sigma_i)]-\sigma_e{}^2(1-r^2)\nabla^2 I(X)/2$$
$$=\sigma_e{}^4(1-r^4)\nabla^4 I(X)/8+o(\sigma_e{}^4),\sigma_e\to 0.$$

1.8. COMPUTING 2-D GRADIENTS OF ARBITRARY ORDER

Let $Q(u,v)$ be a polynomial of total degree q in u, v and $Q(D)\equiv Q(\partial/\partial x,\partial/\partial y)$. The 2-D convolution $I_\sigma(X)=I(X)\otimes G(X,\sigma)$ also has the pointwise simultaneous approximation property of an arbitrary order:

Theorem 1.8.1.

If $I(X)$ is a bounded continuous function, q-differentiable at a point X,

$$\lim_{\sigma\to 0} I(X)\otimes[Q(D)G(X,\sigma)]=Q(D)I(X),$$

the relation holding uniformly in $X\in S\subset R^2$ if $I(X)$ is uniformly q-differentiable in S. Moreover, if $I(X)$ is $2p+q$-differentiable at X, then

$$Q(D)[I(X)\otimes G(X,\sigma)-I(X)]=\sum_k\sigma^{2k}\sum_{\lambda,\mu}((2\lambda)!!(2\mu)!!)^{-1}$$
$$\times\left(\partial^{2k}[Q(D)I(X)]/\partial x^{2\lambda}\partial y^{2\mu}\right)+o(\sigma^{2p}),$$

where the sums run over $2\le k\le p$, $\lambda,\mu\ge 0$, with $\lambda+\mu=k$, and the relation holds uniformly in $X\in S\subset R^2$, if $I(X)$ is uniformly $2p+q$-differentiable in S.

Proof: It is enough to consider the case $Q(D)=\partial^{l+m}/\partial x^l\partial y^m$, where l, $m\ge 0$, and $l+m\le q$. Since q-differentiability implies $l+m$-differentiability,

$$I(X)\otimes Q(D)G(X,\sigma)=\int\!\!\int I(X-H)Q(D)G(H,\sigma)dH$$

$$=\int\!\!\int[I(X)+\sum_{1\le k\le l+m}(k!)^{-1}(-H\cdot\nabla)^k I(X)+\mathrm{o}(H^{l+m})]Q(D)G(H,\sigma)dH$$

$$=(k!)^{-1}l!m![k!/(l!m!)]Q(D)I(X)+\int\!\!\int[\mathrm{o}(|H|^{l+m})]Q(D)G(H,\sigma)dH$$

$$=Q(D)I(X)+\int\!\!\int[\mathrm{o}(|H|^{l+m})]Q(D)G(H,\sigma)dH.$$

As for an arbitrary $\varepsilon>0$, there exists a $\delta>0$, such that

$$\mathrm{o}(H^{l+m}<\left(\varepsilon+H^2/\delta^2\right)H^{l+m},\ H\in \mathrm{R}^2,$$

from the earlier Lemma,

$$\int\!\!\int[\mathrm{o}(H^{l+m})]Q(D)G(H,\sigma)dH$$

$$\le\int\!\!\int\!\!\int(\varepsilon+H^2/\delta^2)H^{l+m}\sum_{r,s\ge0,r-s\le l}c_{rs}^{[l]}(\xi^s/\sigma^r)$$

$$\times\sum_{t,u\ge0,t-u\le m}c_{tu}^{[m]}(\eta^u/\sigma^t)\,\mathrm{G(H,\sigma)dH}$$

$$\le\varepsilon\mathrm{O}\left(\sum_{r,s\ge0,r-s\le l,t,u\ge0,t-u\le m}\sigma^{l+m-(r-s)-(t-u)}\right)$$

$$+\mathrm{O}\left(\sum_{r,s\ge0,r-s\le l,t,u\ge0,t-u\le m}\delta^{-2}\sigma^{2+l+m-(r-s)-(t-u)}\right)$$

$$\le\varepsilon\mathrm{O}(1)+\mathrm{O}(\sigma^2/\delta^2).$$

Hence the o-term goes to zero as $\sigma\to0$, and $I(X)\otimes[Q(D)G(X,\sigma)]\to Q(D)I(X)$ follows. If $I(X)$ is uniformly q-differentiable in S, the $[\varepsilon\mathrm{O}(1)+\mathrm{O}(\sigma^2/\delta^2)]$- term is independent of $\mathrm{X}\in$ S, and the uniformity of the limit relation follows.

If $I(X)$ is $2p$+q-differentiable at X,

$$Q(D)[I(X)\otimes G(X,\sigma)-I(X)]$$

$$=\int\!\!\int\sum_{0\le k\le q+2p}(\mathrm{k!})^{-1}(-H\cdot\nabla)^k I(X)+\mathrm{o}(H^{q+2p})\,Q(D)G(H,\sigma)dH$$

$$=\int\!\!\int\sum_{l+m+1\le k\le q+2p}(\mathrm{k!})^{-1}(-H\cdot\nabla)^k I(X)+\mathrm{o}(H^{q+2p})\,Q(D)G(H,\sigma)dH.$$

As

$$\int\!\!\int\sum_{l+m+1\le k\le q+2p}(k!)^{-1}[(-H\cdot\nabla)^k I(X)]Q(D)G(H,\sigma)dH$$

$$=\sum_{l+m+1\le k\le q+2p}(k!)^{-1}\sum_{2(\lambda+\mu)=k-(l+m)}[(l+2\lambda)!(m+2\mu)!/\{(2\lambda)!!(2\mu)!!\}]\sigma^{k-(l+m)}$$

$$\times[k!/\{(l+2\lambda)!(m+2\mu)!\}]\partial^{2(\mu+\lambda)}Q(D)I(X)/\partial x^{2\lambda}\partial y^{2\mu}$$

$$= \sum_{1\le j\le q+2p-(l+m)} \sigma^{j} \sum_{2(\lambda+\mu)=j} [1/((2\lambda)!!(2\mu)!!)] \partial^{2j} Q(D) I(X) / \partial x^{2\lambda} \partial y^{2\mu} + o(\sigma^{2p})$$

$$= \sum_{1\le j\le p} \sigma^{2j} \sum_{(\lambda+\mu)=j} [(2\lambda)!!(2\mu)!!] \partial^{2j} Q(D) I(X) / \partial x^{2\lambda} \partial y^{2\mu}] + o(\sigma^{2p}),$$

and the $o(H^{q+2p})$-term contributes to $o(\sigma^{2p})$ is clear from:

$$\int_{\cdot\cdot}^{\cdot\cdot} [o(H^{l+m})] Q(D) G(H,\sigma) dH$$

$$\le \int_{\cdot\cdot}^{\cdot\cdot} [(\varepsilon + |H|^2/\delta^2)|H|^{q+2p}] \sum_{r,s\ge 0, r-s\le l} c_{rs}^{[l]} (\xi^s/\sigma^r) \sum_{t,u\ge 0, t-u\le m} c_{tu}^{[m]} (\eta^u/\sigma^t) \times G(H,\sigma) dH$$

$$\le \varepsilon O\left(\sum_{r,s\ge 0, r-s\le l,\ t,u\ge 0, t-u\le m} \sigma^{q+2p-(r-s)-(t-u)}\right) + O\left(\sum_{r,s\ge 0, r-s\le l,\ t,u\ge 0, t-u\le m} \delta^{-2} \sigma^{2+q+2p-(r-s)-(t-u)}\right)$$

$$\le \varepsilon O(\sigma^{2p}) + O(\sigma^{2p+2}/\delta^2),$$

(as $2p+q-(l+m)+l-(r-s)+m-(t-u) \ge 2p+q-(l+m) \ge 0$), the asymptotic relation follows. If $I(X)$ is uniformly $2p+q$-differentiable in S, δ can be chosen to be independent of $X \in S$, so the $[\varepsilon O(\sigma^{2p}) + O(\sigma^{2p+2}/\delta^2)]$ - term are independent of $X \in S$, completing the proof.

1.9. ERROR FOR SMOOTH GRADIENT FUNCTIONS IN PLANE

Here, C denotes the Banach space of all bounded continuous functions on R^2 normed by: $\|I\| = \sup_{X\in R^2} \{|I(X)|\}$, and C^2 denotes the space of bounded 2-differentiable functions $I(X)$ on R^2, having all partial derivatives upto order 2 bounded on R^2. The space C^2 is normed by $\|I\|_{C,2} = \|I\| + \|I\|_2$, where $\|I\|_2 = \max\{\|I_{xx}\|, \|I_{yy}\|, (4/\pi)\|I_{xy}\|\}$ is a semi-norm on C^2 and

$$K(I;t) = \inf_{g\in C^2} \{\|I-g\| + t^2 \|g(X)\|_2\}, I(X) \in C$$

defines the Peetre's K-functional. For $I \in C$,

$$I(X) \otimes G(X,\sigma) - I(X) = \int_{\cdot\cdot}^{\cdot\cdot} [I(X-\sigma H) - I(X)] G(H,1) dH,$$

by Lebesgue's dominated convergence theorem, implies that

$$I(X) \otimes G(X,\sigma) \to I(X), \text{ as } \sigma \to 0.$$

This convergence is uniform on compact subsets of R^2. If $I(X)$ is uniformly continuous on R^2,

$$\omega(\delta) = \sup\{ I(X+H) - I(X) : H \le \delta \} \to 0, \delta \to 0,$$

and then the convergence $I(X) \otimes G(X,\sigma) \to I(X), \sigma \to 0,$ is uniform on R^2.

Lemma 1.9.1.

$$I(X) \otimes G(X,\sigma) - I(X) \le (3/2)\sigma^2 \, I_{\,2}, I \in C^2.$$

Proof: If $I(X) \in C^2$, by mean value theorem 0

$$I(X+H) - I(X) - H \cdot \nabla I(X) \le (1/2)[h_1^{\,2} I_{xx} + h_2^{\,2} I_{yy} + 2h_1 h_2 I_{xy}](X + \theta H),$$

$0 < \theta < 1$, so that

$$\begin{aligned} I(X) \otimes G(X,\sigma) - I(X) &\le (1/2) \iint [h_1^{\,2} I_{xx} + h_2^{\,2} I_{yy} + 2 h_1 \, h_2 \, I_{xy}] G(H,\sigma) dH \\ &\le (I_{\,2}/2) \iint [h_1^{\,2} + h_2^{\,2} + (\pi/2)(h_1 \, h_2)] G(H,\sigma) dH \\ &= (3/2) \, I_{\,2} \sigma^2. \end{aligned}$$

Lemma 1.9.2.

$$I(X) \otimes G(X,\sigma)_{\,2} \le 2\sigma^{-2} \, I, I \in C.$$

Proof:
$$\begin{aligned} [I(X) \otimes G(X,\sigma)]_{xx} &= \iint I(X-H) G_{xx}(H,\sigma) dH \\ &= \iint [I(X-H)\{-1/\sigma^2 + (h_1/\sigma^2)^2\} G(H,\sigma)] dH \\ &\le 2 \, I| \sigma^{-2}. \end{aligned}$$

Similarly,

$$[I(X) \otimes G(X,\sigma)]_{yy} \le 2 \, I \, \sigma^{-2}.$$

Also,

$$\begin{aligned} [I(X) \otimes G(X,\sigma)]_{xy} &= \iint I(X-H) G_{xy}(H,\sigma) dH \\ &= \iint [I(X-H)\{xy/\sigma^4\} G(H,\sigma)] dH \\ &\le (2/\pi) \, I \, \sigma^{-2}. \end{aligned}$$

Hence, $I(X)\otimes G(X,\sigma)\ _2 \le \max\{2\ I|\sigma^{-2}, (8/\pi^2)\|I|\sigma^{-2}\} = 2\sigma^{-2}\ |I\|$, completing the proof of the Lemma.

Lemma 1.9.3.

$\|I(X)\otimes G(X,\sigma) - I(X)\ \le 2K(I;\sigma), I \in C.$

Proof: Let $I(X)\in C$. Then, for $g\in C^2$, we have:

$$\begin{aligned} I(X)\otimes G(X,\sigma)-I(X) &\le I(X)-g(X) + g(X)-g(X)\otimes G(X,\sigma) \\ &\quad + [g(X)-I(X)]\otimes G(X,\sigma) \\ &\le 2\ I(X)-g(X)\ + (3/2)\sigma^2\ g\ _2. \end{aligned}$$

Taking infimum over $g\in C^2$, we have $I(X)\otimes G(X,\sigma)-I(X)\ \le 2K(I;\sigma)$, completing the proof.

Theorem 1.9.4.

Let $I(X)\in C$ and $\psi\in S$. Then:

(i) $I(X)\otimes G(X,\sigma)-I(X)\ = O(\psi(\sigma^2)), \sigma\to 0$, iff $K(I;t) = O(\psi(t^2))$, $t\to 0$;

(ii) $|I(X)\otimes G(X,\sigma)-I(X)| = o(\psi(\sigma^2)), \sigma\to 0$, iff $K(I;t) = o(\psi(t^2)), t\to 0$; and,

(iii) $I(X)\otimes G(X,\sigma)-I(X)\ \sim K(I;\sigma)$, if $K(I;t^{1/2})\in S$.

Proof: $I(X)\otimes G(X,\sigma)-I(X)\ = O(\psi(\sigma^2))$, if $K(I;t) = O(\psi(t^2))$, and $I(X)\otimes G(X,\sigma)-I(X)| = o(\psi(\sigma^2))$, if $K(I;t) = o(\psi(t^2))$, by the previous Lemma. If $I(X)\otimes G(X,\sigma)-I(X)\ = O(\psi(\sigma^2))$, then

$$\begin{aligned} K(I;t) &\le K([I(X)-I(X)\otimes G(X,\sigma)];t) + K([I(X)-g(X)]\otimes G(X,\sigma);t) \\ &\quad + K(g(X)\otimes G(X,\sigma);t) \\ &\le O(\psi(\sigma^2)) + 2(t/\sigma)^2\ I(X)-g(X)\ + t^2\ g\ _2. \end{aligned}$$

Hence, taking infimum over $g\in C^2$, $K(I;t)\le O(\psi(\sigma^2)) + 2(t/\sigma)^2 K(I;\sigma)$, and the rest of the results follow from Lemma 1.2.2.

Corollary 1.9.5.

Let $I \in C$, and $0 < \alpha < 2$. Then:

(i) $I(X) \otimes G(X,\sigma) - I(X) = O(\sigma^{\alpha}), \sigma \to 0$, iff $K(I(X);t) = O(t^{\alpha}), t \to 0$;

(ii) $I(X) \otimes G(X,\sigma) - I(X) = o(\sigma^{\alpha}), \sigma \to 0$, iff $K(I(X);t) = o(t^{\alpha}), t \to 0$; and,

(iii) $\| I(X) \otimes G(X,\sigma) - I(X) \| \sim K(I;\sigma)$, if $K(I(X);t) \sim t^{\alpha}$.

With $Q(u,v)$, as before, a polynomial in u, v, and $Q(D) \equiv Q(\partial/\partial x, \partial/\partial y)$, the previous Theorem and the Corollary imply the following:

Theorem 1.9.6.

Let $Q(D)I(X) \in C$, and $\psi \in S$. Then:

(i) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = O(\psi(\sigma^2)), \sigma \to 0$,

iff $K(Q(D)I(X);t) = O(\psi(t^2)), t \to 0$;

(ii) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = o(\psi(\sigma^2)), \sigma \to 0$,

iff $K(Q(D)I(X);t) = o(\psi(t^2)), t \to 0$; and,

(iii) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) \sim K(I;\sigma)$, if $K(Q(D)I(X);t^{1/2}) \in S$.

Corollary 1.9.7.

Let $Q(D)I(X) \in C$, and $0 < \alpha < 2$. Then:

(i) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = O(\sigma^{\alpha}), \sigma \to 0$,

iff $K(Q(D)I(X);t) = O(t^{\alpha}), t \to 0$;

(ii) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) = o(\sigma^{\alpha}), \sigma \to 0$,

iff $K(Q(D)I(X);t) = o(t^{\alpha}), t \to 0$; and,

(iv) $I(X) \otimes [Q(D)G(X,\sigma)] - Q(D)I(X) \sim K(I;\sigma)$, if $K(Q(D)I(X);t) \sim t^{\alpha}$.

Finally, we complete the 2-D saturation theory for the convolution approximation with the Gaussian filter. To this purpose, let the circular average $I_\rho(X)$ of a function $I(X) \in C$ be defined as:

$$I_\rho(X) = \int_{C(X,\rho)} I(T)ds(T),$$

where $C(X,\rho)$ is the circle with centre X and radius ρ, T represents a general point, and, $ds(T)$ is the arc length element. We shall call $I(X)$ harmonic $I_\rho(X) = I(X)$, $X \in R^3$, $\rho > 0$.

Lemma 1.9.8.

If $I(X)$ possesses continuous second order partial derivatives on R^2, and, if $\nabla^2 I(X) \in C$, $I_\rho(X) - I(X) \le \rho^2 \nabla^2 I(x)/4$, $\rho > 0$, where equality holds if $I(X) \equiv |X|^2$.

Proof:

$$\begin{aligned} I_\rho(X) - I(X) &= (2\pi)^{-1} \int_{|T|=1} [I(X - \rho T) - I(X)]ds \\ &= (2\pi)^{-1} \int_{[0,\rho]} \left(\int_{|T|=1} [\partial I(X - rT)/\partial n]ds \right) dr \\ &= (2\pi)^{-1} \int_{[0,\rho]} \left(r^{-1} \int_{|T-X|=r} [\nabla I(T) \cdot n]ds \right) dr. \end{aligned}$$

Hence, by the Green's theorem in the plane,

$$\begin{aligned} I_\rho(X) - I(X) &= (2\pi)^{-1} \int_{[0,\rho]} \left(r^{-1} \iint_{|T-X|\le r} div(\nabla I(T))dxdy \right) dr \\ &= (2\pi)^{-1} \int_{[0,\rho]} \left(r^{-1} \iint_{|T-X|\le r} [\nabla^2 I(T)]dxdy \right) dr, \end{aligned}$$

so that

$$I_\rho(X) - I(X) \le \left\{ (2\pi)^{-1} \int_{[0,\rho]} \left(r^{-1} \iint_{|T-X|\le r} dxdy \right) dr \right\} \nabla^2 I = \rho^2 \nabla^2 I / 4.$$

Theorem 1.9.9.

Let $I(X) \in C$. Then:

(iii) $|I(X) \otimes G(X,\sigma) - I(X)| = O(\sigma^2), \sigma \to 0$, iff $|I_r(X) - I(X)| = O(r^2), r \to 0$;

and,

(iv) $I(X)\otimes G(X,\sigma)-I(X)=o(\sigma^2),\sigma\to 0$, iff $I(X)$ is harmonic;

As an immediate consequence of this Theorem, we have the follow

Corollary 1.9.10.

Let $Q(D)I(X)\in C$. Then:

(iii) $I(X)\otimes[Q(D)G(X,\sigma)]-Q(D)I(X)=O(\sigma^2),\sigma\to 0,$

iff $Q(D)I_r(X)-Q(D)I(X)=O(r^2),r\to 0$; and,

(iv) $I(X)\otimes[Q(D)G(X,\sigma)]-Q(D)I(X)=o(\sigma^2),\sigma\to 0,$

iff $Q(D)I(X)$ is harmonic.

Proof of the Theorem:

If $I_r(X)-I(X)=O(r^2),r\to 0$, as $I(X)\in C$, $I_r(X)-I(X)\le Mr^2,r>0$, for an $M>0$. So,

$$\begin{aligned}I(X)\otimes G(X,\sigma)-I(X)&=\int\!\!\int[I(X-T)-I(X)]G(T,\sigma)dT\\&=\int_{[0,\infty)}\left(\int_{C(0,\rho)}[I(X-T)-I(X)]G(T,\sigma)ds(T)\right)d\rho\\&=\int_{[0,\infty)}2\pi\rho[I_\rho(X)-I(X)](2\pi\sigma^2)^{-1}\exp\{-\rho^2/2\sigma^2\}d\rho,\end{aligned}$$

whence

$$\begin{aligned}I(X)\otimes G(X,\sigma)-I(X)&\le(2\pi\sigma^2)^{-1/2}\pi M\int_{[0,\infty)}2\rho^3(2\pi\sigma^2)^{-1/2}\exp\{-\rho^2/2\sigma^2\}d\rho\\&=2M\sigma^2=O(\sigma^2).\end{aligned}$$

Conversely, if $I(X)\otimes G(X,\sigma)-I(X)=O(\sigma^2)$, writing:

$$I^{[\tau]}(X)=I(X)\otimes G(X,\tau),$$

we have, $I^{[\tau]}(X)\otimes G(X,\sigma)-I^{[\tau]}(X)=O(\sigma^2)$. Hence, for a $B>0$,

$$I^{[\tau]}(X)\otimes G(X,\sigma)-I^{[\tau]}(X)\le B\sigma^2,$$

for all σ sufficiently small. By

$$I^{[\tau]}(X)\otimes G(X,\sigma)-I^{[\tau]}(X)=(\sigma^2/2)\nabla^2 I^{[\tau]}(X)+o(\sigma^2),$$

we have, $\nabla^2 I^{[\tau]}(X) \le 2B$, and by the previous Lemma,

$$(I^{[\tau]})_\rho(X) - I^{[\tau]}(X) \le \rho^2 \nabla^2 I^{[\tau]}(X) \Big/ 4 \le B\rho^2/2.$$

Taking limit as $\tau \to 0$, $I_\rho(X) - I(X) \le B\rho^2/2 = \mathrm{O}(\rho^2)$, which proves part (i). For (ii), if $I(X)$ is harmonic, $0 = \|I(X) \otimes G(X,\sigma) - I(X)\| = \mathrm{o}(\sigma^2)$. Conversely, from the latter, $I^{[\tau]}(X) \otimes G(X,\sigma) - I^{[\tau]}(X) = \mathrm{o}(\sigma^2)$, uniformly in τ. Hence, given any $\varepsilon > 0$, there exists a $\sigma_0 > 0$, such that

$$\left| I^{[\tau]}(X) \otimes G(X,\sigma) - I^{[\tau]}(X) \right| \le \varepsilon\sigma^2,\ 0 < \sigma < \sigma_0.$$

It follows that: $\nabla^2 I^{[\tau]}(X) \le 2\varepsilon$, and

$$(I^{[\tau]})_\rho(X) - I^{[\tau]}(X) \le \varepsilon\rho^2/3.$$

Letting $\tau \to 0$, $I_\rho(X) - I(X) \le \varepsilon\rho^2/3, \rho > 0$. As ε is arbitrary $I_\rho(X) \equiv I(X)$, completing the proof.

1.10. NUMERICAL SIMULATONS AND DISCUSSION

In this secton we present some numerical simulation results for the basic Gaussian filters with σ^{-1} = 128, 64, 32, 16, and 8.

In Figure 1.1 (Original T_2 weighted and corresponding convolved images with Gaussian filter), the first row first image is the original cross section, followed by its Gaussian filtered version with σ^{-1} = 128. The second and the third row images correspond to the Gaussian filter with σ^{-1} = 64, 32, 16, and 8, respectively in the natural order. The filtered image with σ^{-1} = 128 gives quite a good approximation and does not seem to be visibly any different from the original image. The case σ^{-1} = 64 seems to induce sufficient smoothness without blurring any important structures. From σ^{-1} = 32 downwards the increasing blurring is apparent which outlines only grosser regions and hides the finer features.

The percentage relative L1 and L2 errors in the approximations for the Gaussian filter with σ^{-1} = 128, 64, 32, 16, and 8 are summarized in the Table 1.1, which are defined as:

$$\text{Percentage Relative L1 error} = \frac{\sum_{\text{pixels}} f(i) - \tilde{f}(i)}{\sum_{\text{pixels}} f(i)} \times 100,$$

$$\text{Percentage Relative L2 error} = \left(\frac{\sum_{\text{pixels}} f(i) - \tilde{f}(i)^2}{\left(\sum_{\text{pixels}} |f(i)|^2 \right)^{1/2}} \right)^{1/2} \times 100,$$

where $f(i)$ and $\tilde{f}(i)$ represent original and filtered images.

σ^{-1}	128	64	32	16	8
% Relative L1 Error	0.19906	4.69717	9.52364	14.63921	17.40059
% Relative L2 Error	0.17313	4.23735	9.62608	15.87283	19.08742

Table 1.1: % Relative Errors in the Approximation of the Original Cross-Section with Gaussian Filter

Figure 1.2 (MLE Segmented images of Convolutions with Gaussian filter) contains MLE-based segmentations associated with the images in Figure 1.1. The number of iterations taken for original image is 59400 and for σ^{-1} = 128, 64, 32, 16, 8 are 18600, 13550, 3050, 7750, 3700 respectively. The segmentations associated with the original and σ^{-1} = 128 are unnecessarily detailed. The cases corresponding to σ^{-1} = 64, and σ^{-1} = 32 seem to be of practical use, while those with σ^{-1} = 16, and σ^{-1} = 8 provide grosser simplifications.

For σ^{-1} = 32 and 16, there is a splitting of edema into three regions, whereas these merge to give two regions for σ^{-1} = 8 and boundary of edema thickens with the decreasein the value of σ^{-1}. All separate regions corresponding to edema are represented as different stages of its growth inside the brain parenchyma. As filtering increases, the segments within cranial part merge to give one segment and there is a formation of boundary between the major segments comprising of brain and its cranial part. Excessive filtering is useful in the separation of brain from its cranial part.

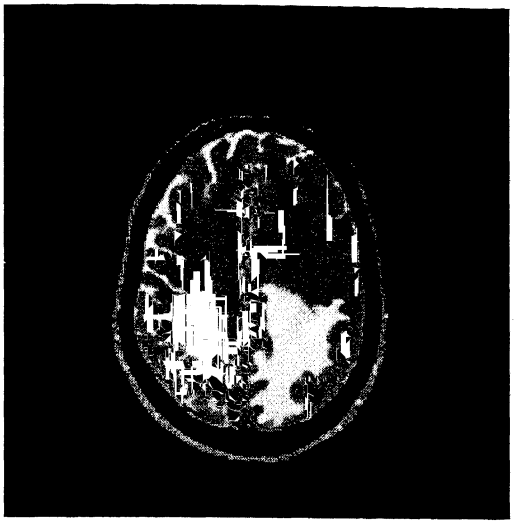
In conclusion, for the Gaussian filter the choice σ^{-1} = 64, 32 seem to be satisfactory for the purpose of segmentation, and the MLE based segmentation appears to be promising.

Figure 1.3 (Gaussian Difference Magnitudes) presents the original cross section of Figure 1.1 and the magnitudes of the differences of the each of the rest of the images in Figure 1.1 with the previous one. As these approximate the Laplacian, it is clear that the original image possesses non-uniform graininess and that the segments of interest do not have uniform intensity along and across the boundaries. As expected, the Laplacian boundaries thicken with decreasing σ^{-1}. It may be concluded that the Laplacian in the present case does not seem to provide closed and continuous boundaries. Since filters for the computation of the Laplacian, and the DOG filter behave similarly to the absolute differences, their utility for the MRI images of the present category may not be very competitive with the MLE based segmentation.

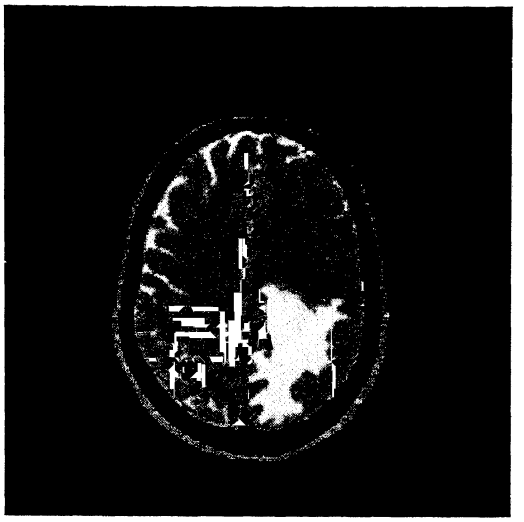
Figure 1.4 (Sobel maps (Gaussian filter)) displays the Sobel maps of the images in Figure 1.1. They possess the same general behaviour with respect to σ as the difference magnitudes. However, they provide a much better description of the edges as compared with the difference magnitudes. The discontinuities and non-closedness are observable.

Figure 1.5 (Mathematical Edges of MLE-segmented original and filtered images) displays the precise edges from the respective MLE-segmented images. Figure 1.6 (Edges traced from original and Gaussian filtered images) displays the results of using a certain popular edge tracing commercial program on the filtered images. Figure 1.7 (Edges traced from MLE-segmented original and filtered images) does the same on the segmented images.

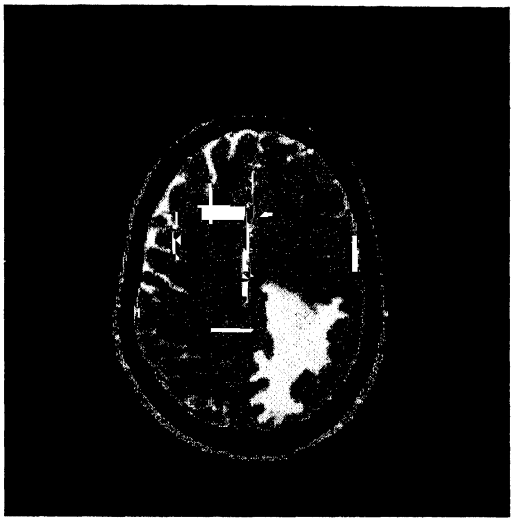
Results of Marr-Hildreth operator with different values of σ (1.0, 1.61, 2.0, 2.576, 3.0, 4.0) are shown in Figure 1.8. It is observed that the obtained closed contours are well represented in the cases of $\sigma = 2.0$, 2.576 and 3.0. Figure 1.9 shows the results of Difference-of-Gaussian in two-dimension (A. Huertas and G. Medioni [1986]) with the ratio of 1.6 between the values of $\sigma_e = 0.805\sigma$ and $\sigma_i = 1.288\sigma$, the different values of σ are 1.0, 1.2, 1.4, 1.6, 1.8, 2.0. The number of closed contours decreases with the increase in the value of σ, i.e., due to increase in the smoothness within an image, which is expected.



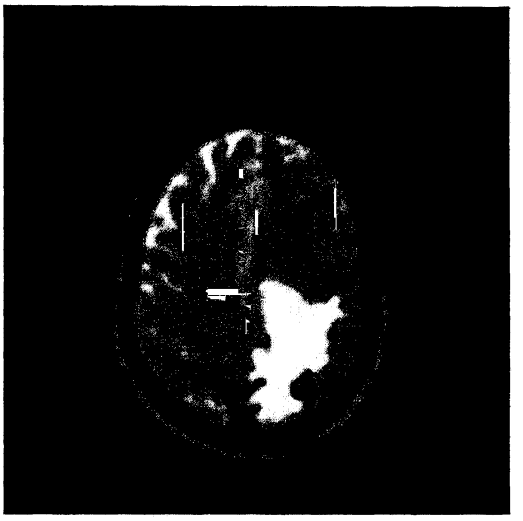
Original



σ^{-1}=128



σ^{-1}=64

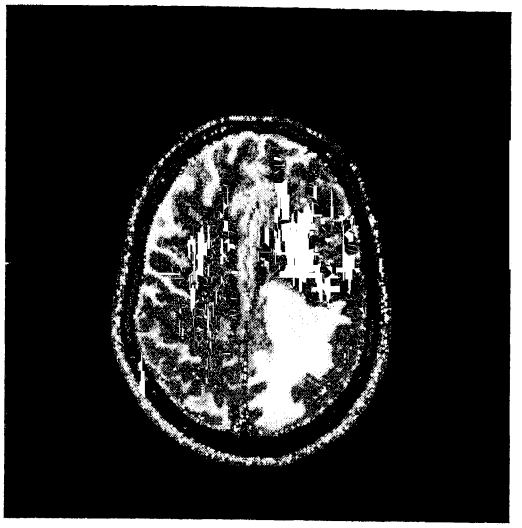


σ^{-1}=32

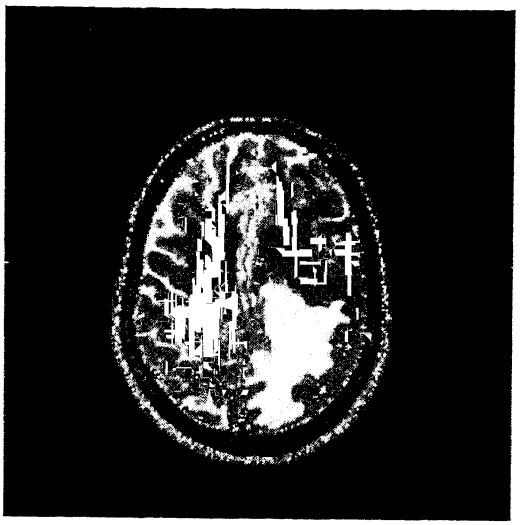
σ^{-1}=16

σ^{-1}=8

Figure 1.1: Original T_2 weighted and Corresponding Filtered Images with Gaussian Filter



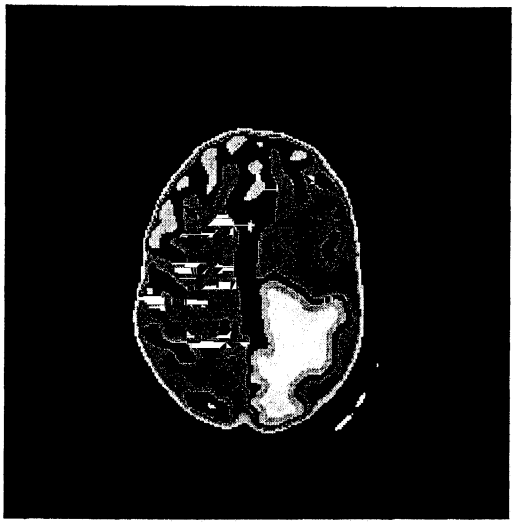
Original



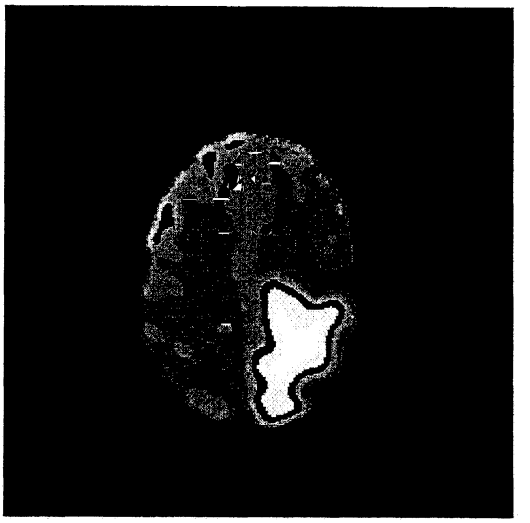
σ^{-1}=128

σ^{-1}=64

σ^{-1}=32

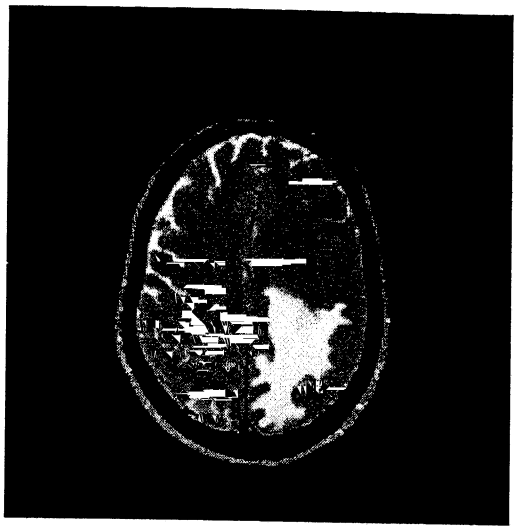


σ^{-1}=16

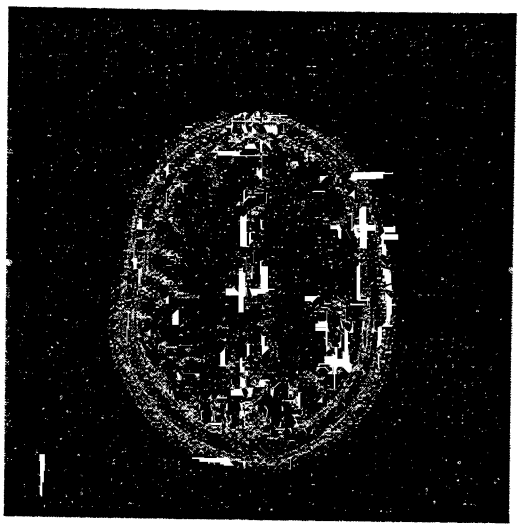


σ^{-1}=8

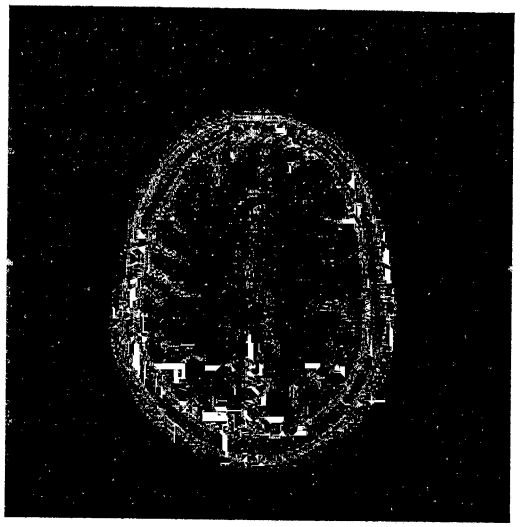
Figure 1.2: MLE Segmented Images of Convolutions with Gaussian Filter



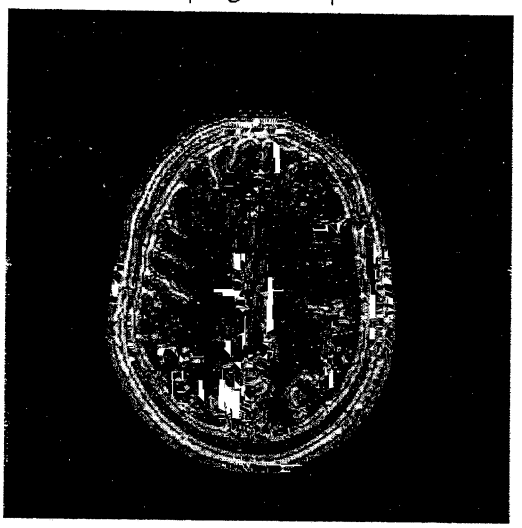
Original



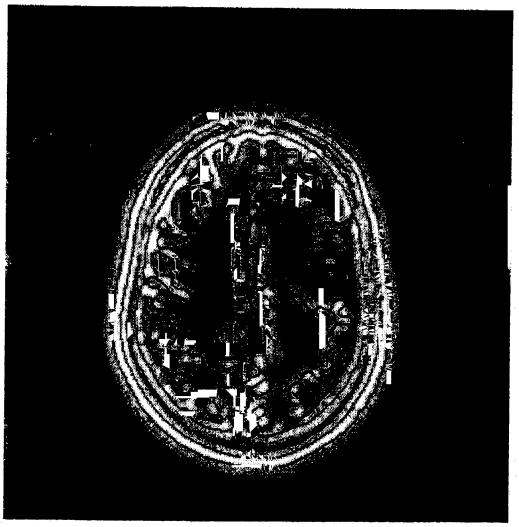
|Original-128|



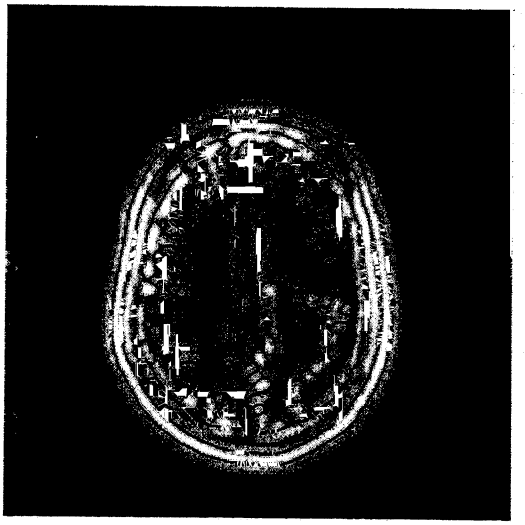
|128-64|



|64-32|

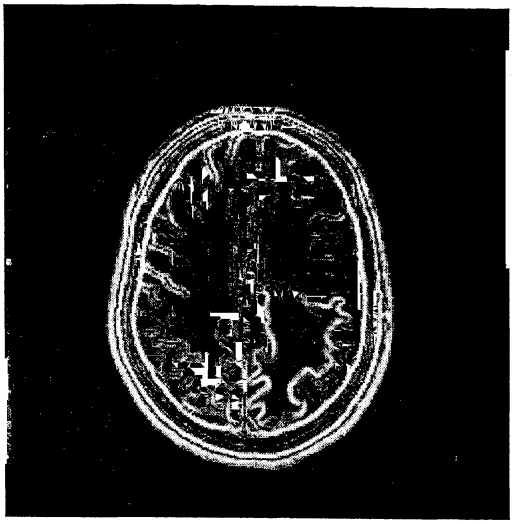


|32-16|



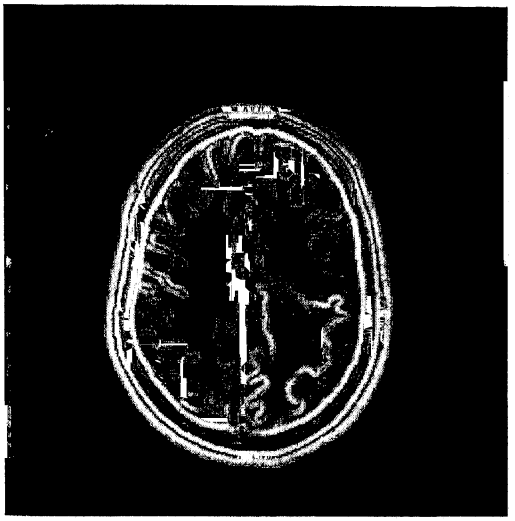
|16-8|

Figure 1.3: Gaussian Difference Magnitudes



Original

σ^{-1}=128



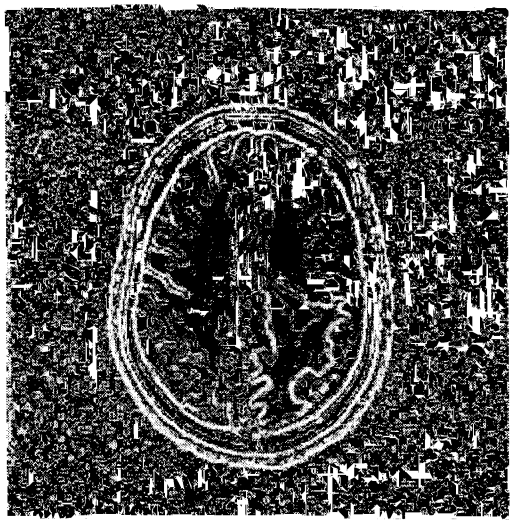
σ^{-1}=64

σ^{-1}=32

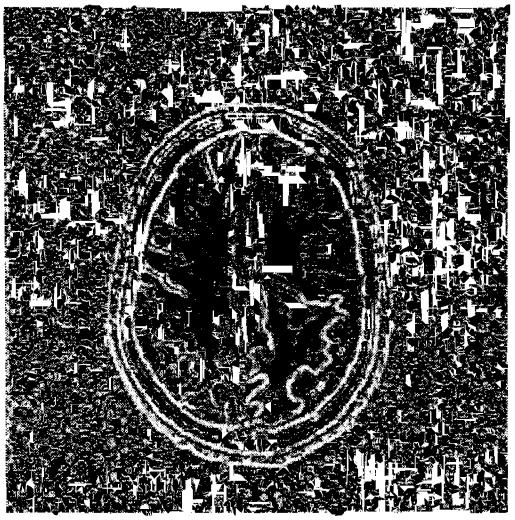
σ^{-1}=16

σ^{-1}=8

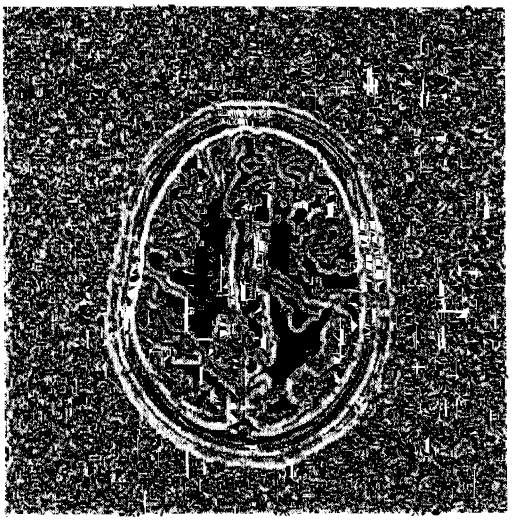
Figure 1.4: Sobel Maps (Gaussian Filter)



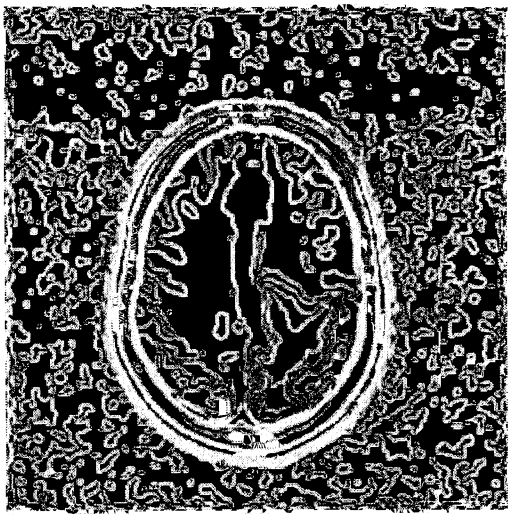
Original



σ^{-1}=128

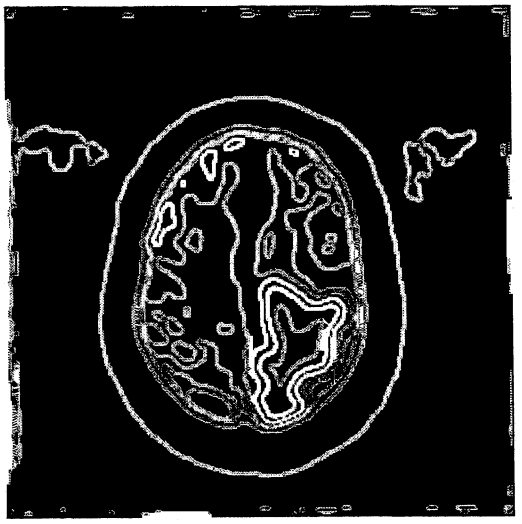


σ^{-1}=64



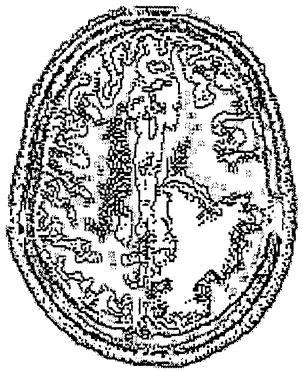
σ^{-1}=32

σ^{-1}=16

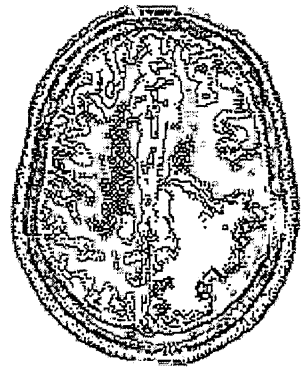


σ^{-1}=8

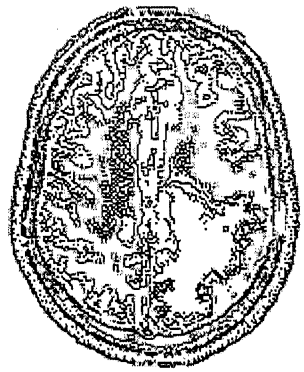
Figure 1.5: Mathematical Edges of MLE Segmented Original and Filtered Images



Original

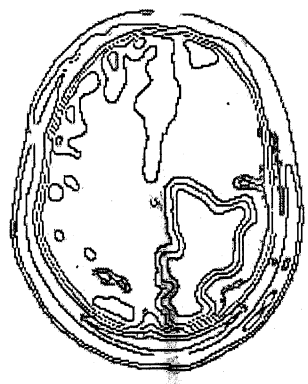


σ^{-1}=128

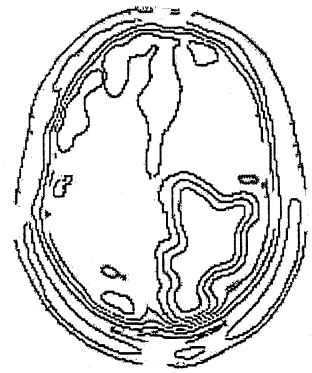


σ^{-1}=64

σ^{-1}=32

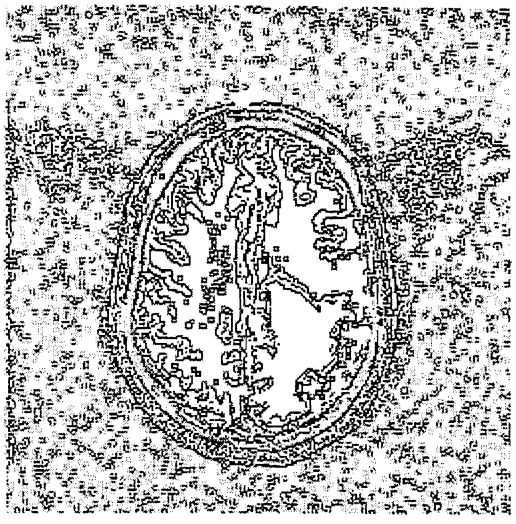


σ^{-1}=16

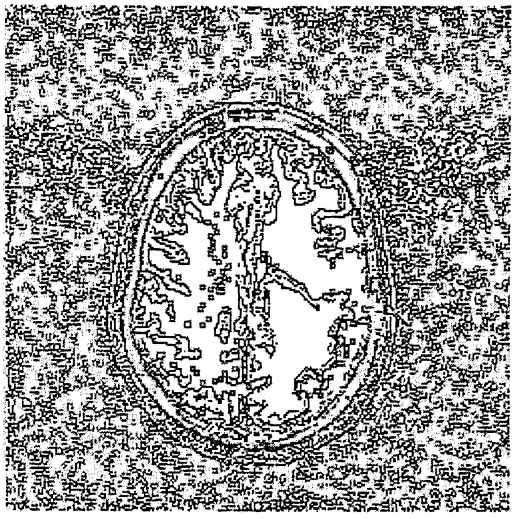


σ^{-1}=8

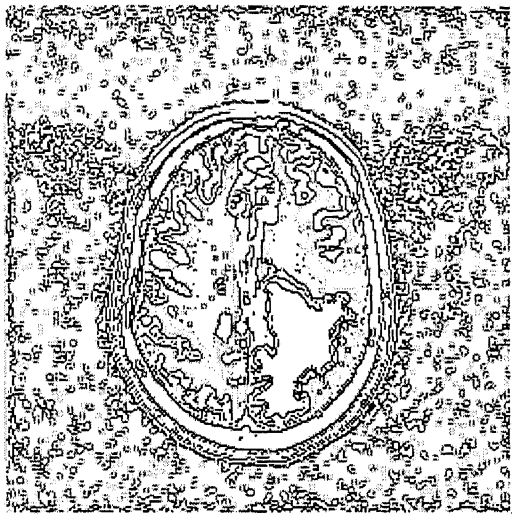
Figure 1.6: Edges Traced from Original and Gaussian Filtered Images



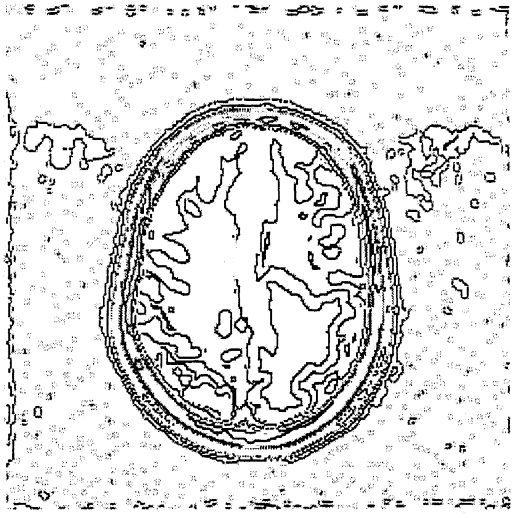
Original



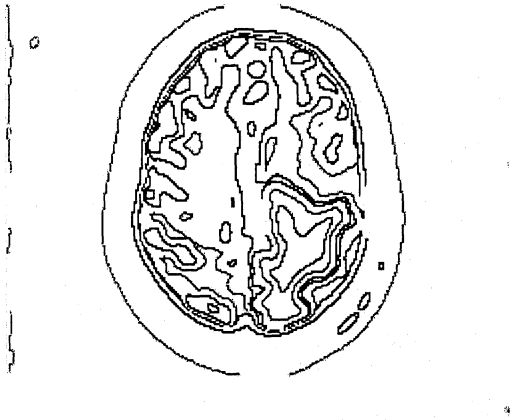
σ^{-1}=128



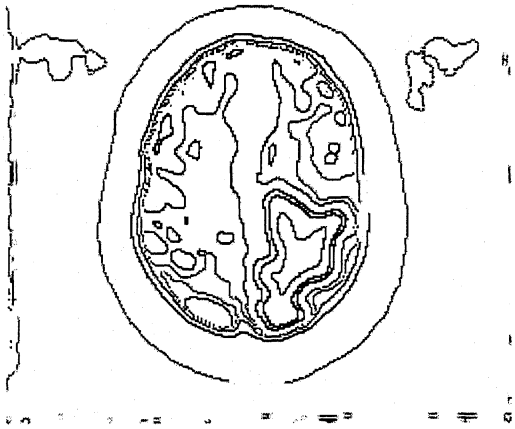
σ^{-1}=64



σ^{-1}=32

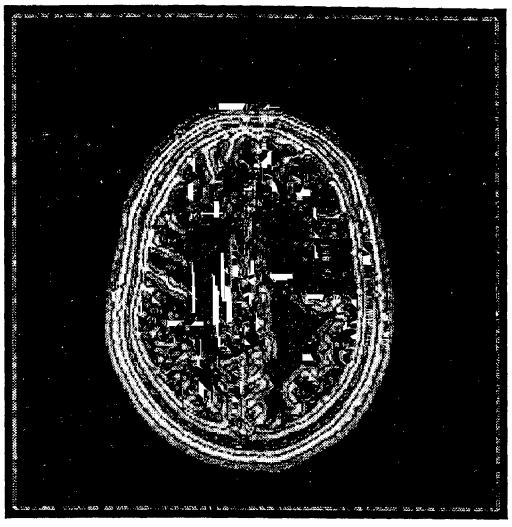


σ^{-1}=16

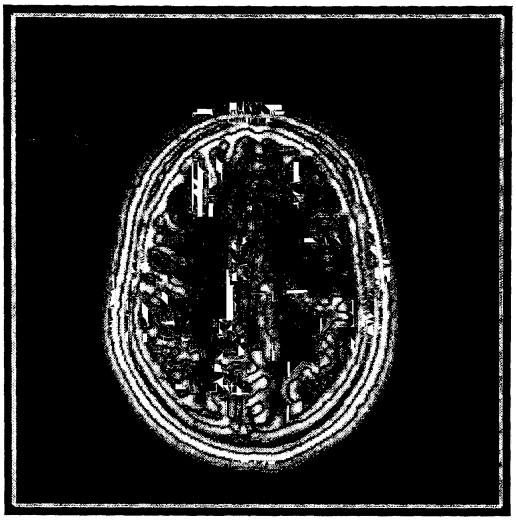


σ^{-1}=8

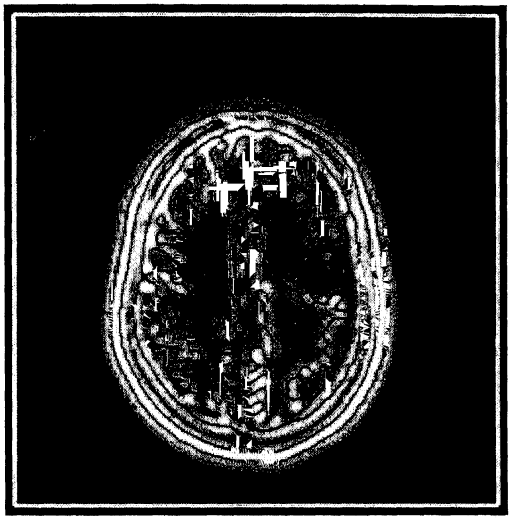
Figure 1.7: Edges Traced from MLE Segmented Original and Filtered Images



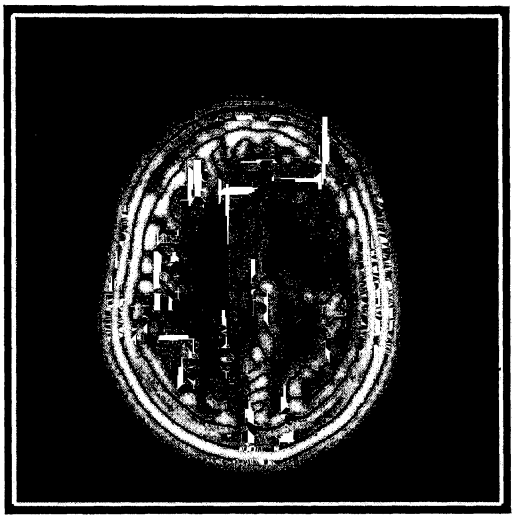
σ=1.0



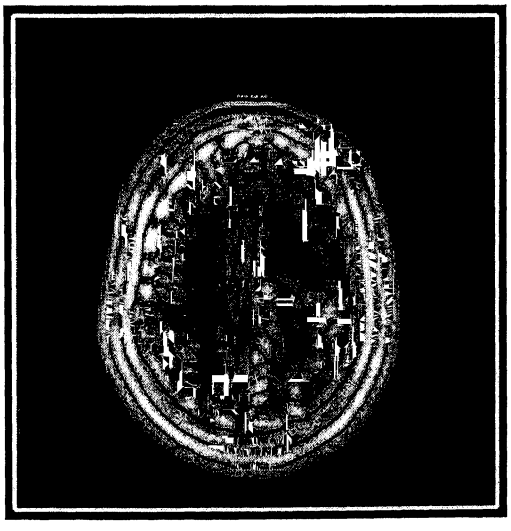
σ=1.6



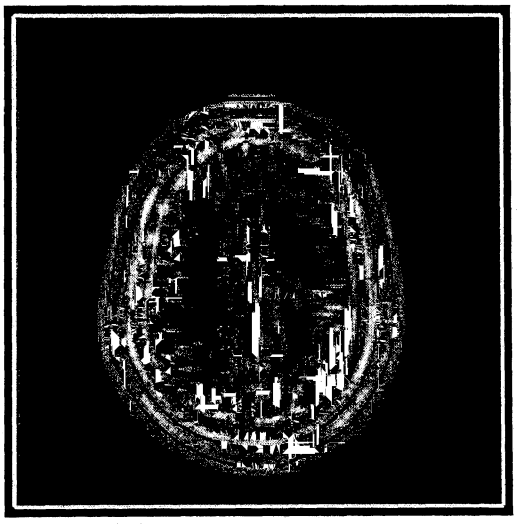
σ=2.0



σ=2.576

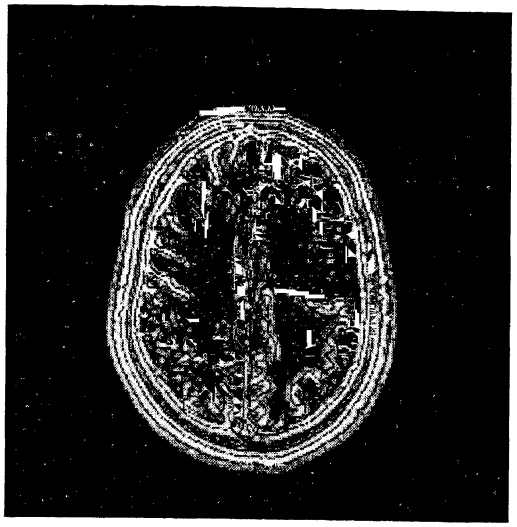


σ=30

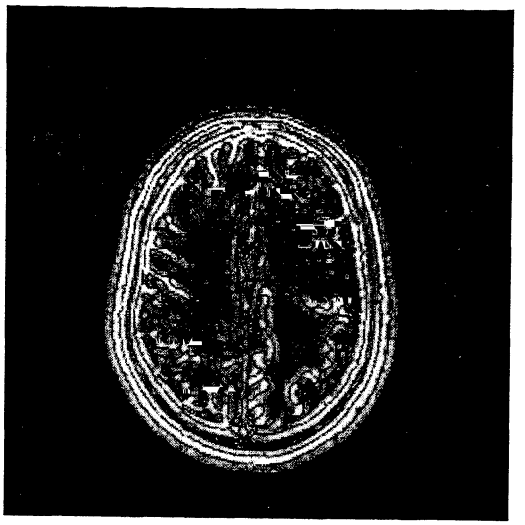


σ=4.0

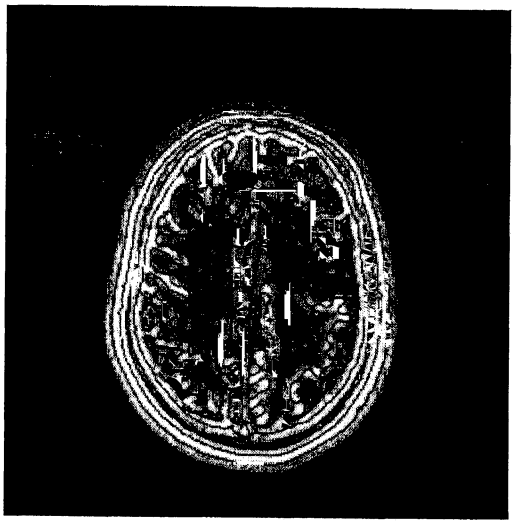
Figure 1.8: Edges via Marr-Hildreth Operator for Different Values of σ



σ=1.0



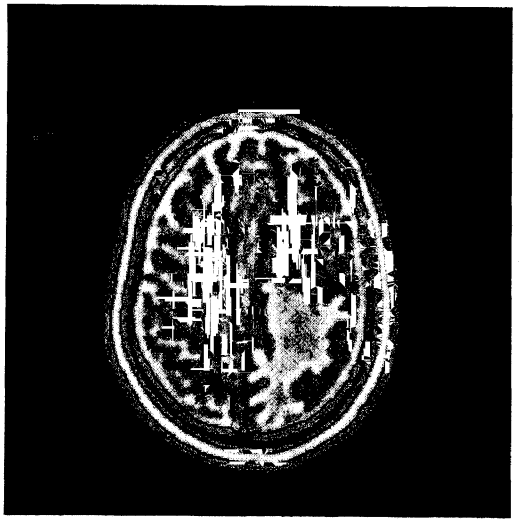
σ=1.2



σ=1.4

σ=1.6

σ=1.8



σ=2.0

Figure 1.9: Edges via Difference-of-Gaussian Operator for Different Values of σ ($\sigma_e = 0805\sigma, \sigma_i = 1.288\sigma$)

CHAPTER II : DIFFERENTIAL SMOOTHING AND BUTTERWORTH FILTERS

In this chapter we study Butterworth and some related filters and specially the degree of approximation of function $f(x)$ by trigonometric polynomial filter

$$T_n(x) = T_n(f;x) = a_0(f)/2 + \sum_{1\le k\le n} \left(a_k(f)\cos kx + b_k(f)\sin kx\right)/\left(1+\theta(k/n)^{2m}\right),$$

where $a_k(f)$ and $b_k(f)$ are the Fourier coefficients of f, which is characterized by a certain minimization property. Section 1.1 is an introduction to simultaneous approximations with a minimization of derivatives of different order. The approximation properties have been derived in section 2.2. Section 2.3 is a study of 1-D Butterworth filters along with its generalization to 2-D, and the 3-D Butterworth filters have been studied in section 2.4. Numerical simulations corresponding to the approximations and the resulting segmentations by 2-D Butterworth filters have been discussed in section 2.5.

2.1. APPROXIMATIONS ALONG WITH DIFFERENT ORDER DERIVATIVE MINIMIZATION

The behavior of the derivatives of approximations to functions representing images could be controlled to quite an extent. The idea is to retain a good degree of approximation while controlling the local variations through minimizing the magnitude of different orders of derivatives of the functions. This is expected to yield better images for the purposes of segmentation as it is well known that the segmentation methods do not perform nicely for cross-sections with a high amount of local intensity changes. Since the intensity variation is governed by derivatives of different orders, it seems natural to try smoothing, without much compromise in the approximation of the cross section, while simultaneously minimizing the derivatives. Let T_n denote a complex trigonometric polynomial of order n. Since the Bernstein

inequality implies that the sequence of functions $\{n^{-m}T_n^{(m)}\}_{m\geq 1}$ is bounded both in $C_{2\pi}$ and ${L_{2\pi}}^2$, by T_n, and as the MRI reconstructions are trigonometric polynomials, at least in the x and y-directions, one could derive meaningful separable filters by minimizing

$$Q_m = (2\pi)^{-1} \int_{[-\pi,\pi]} \left\{ |f(x) - T_n(x)|^2 + \theta\ n^{-m} T_n^{(m)}(x)^2 \right\} dx, \qquad (\theta > 0, f \in {L_{2\pi}}^2),$$

where $f(x)$ represents a complex cross section along rows or columns. Here ${L_{2\pi}}^2$ is the space of 2π-periodic square integrable functions

$$f(x) = a_0/2 + \sum_{k\geq 1} (a_k \cos kx + b_k \sin kx),$$

with the norm given by

$$\|f\| = \left(\int_{[-\pi,\pi]} f(x)^2 dx \right)^{1/2} = \sqrt{\pi} \left\{ {a_0}^2/2 + \sum_{k\geq 1} \left({a_k}^2 + {b_k}^2 \right) \right\}^{1/2},$$

and θ is a parameter. If θ is large the smoothing is expected to get increased. Moreover, by increasing m the smoothing effect ought to decrease, as higher derivatives have only secondary or finer effects on intensity change in digital images (D. Luo [1998]).

2.2. THE MINIMIZING DIFFERENTIAL FILTER

Let T_n denote the space of trigonometric polynomials of order n, $n \geq 0$. The characterization of the form of the trigonometric polynomial T_n minimizing the quantity Q_m of the previous section is given in the following result. The form of the minimizing polynomial makes it depend linearly on the function f. Subsequently the approximation properties of T_n have been derived.

Theorem 2.2.1.

For $f \in {L_{2\pi}}^2, m = 1,2,3,\ldots,$ and $\theta > 0$, the trigonometric polynomial $T_n(x) \in \mathsf{T}_n$ minimizing

$$Q_m = (2\pi)^{-1} \int_{[-\pi,\pi]} \left\{ f(x) - T_n(x)^2 + \theta\, n^{-m} T_n^{(m)}(x)^2 \right\} dx,$$

is unique and is given by

$$T_n(x) = T_n(f;x) = a_0(f)/2 + \sum_{1\le k\le n} \left(a_k(f)\cos kx + b_k(f)\sin kx\right)/\left(1+\theta(k/n)^{2m}\right),$$

where $a_k(f)$, $b_k(f)$ are the Fourier coefficients of the function $f(x)$. The map $f \to T_n$ is linear, and, $\|T_n - f\| \to 0$, $n \to \infty$.

Proof: To find the minimizing T_n, writing

$$f(x) = \sum_{k\ge 0} c_k e^{ikx}, \text{ and } T_n(x) = \sum_{|k|\le n} g_k e^{ikx},$$

and using the orthogonality of the functions e^{ikx},

$$Q_m = \sum_{|k|\le n} \left(|c_k - g_k|^2 + \theta(k/n)^{2m} |g_k|^2 \right) + \sum_{|k|>n} |c_k|^2.$$

The term $\sum_{|k|>n} |c_k|^2$ does not affect the minimization, and it is clear that each of the summands $\left(|c_k - g_k|^2 + \theta(k/n)^{2m} |g_k|^2 \right)$ must separately get minimized. If $c_k = 0$, $g_k = 0$ is the minimizer. Hence with $g_k = \phi\, c_k$, one has to minimize

$$g(\phi) = 1 - \phi^2 + \theta(k/n)^{2m} \phi^2,$$

for which the minimizing ϕ is real and is given by:

$$0 = g'(\phi) = 2\left(\phi\left\{1+\theta(k/n)^{2m}\right\} - 1\right),$$

so that

$$\phi = \left[1+\theta(k/n)^{2m}\right]^{-1},$$

and as

$$g''(\phi) = 2\left(1+\theta(k/n)^{2m}\right) > 0,$$

the minimizer is unique.

Passing to the trigonometric polynomial form, it follows that the trigonometric polynomial $T_n(x)$ minimizing Q_m is given as:

$$T_n(x) = T_n(f;x) = a_0(f)/2 + \sum_{1\le k\le n} \left(a_k(f)\cos kx + b_k(f)\sin kx\right)/\left(1+\theta(k/n)^{2m}\right),$$

with

$$a_k(f)=\pi^{-1}\int_{[-\pi,\pi]} f(x)\cos kx dx,\quad b_k(f)=\pi^{-1}\int_{[-\pi,\pi]} f(x)\sin kx dx,$$

being the Fourier coefficients of $f(x)$. Finally, for the convergence,

$$\begin{aligned} f-T_n &= 2\pi\left\{\sum_{1\le k\le n}\left(\theta(k/n)^{2m}/\left(1+\theta(k/n)^{2m}\right)\right)g_k^2+\sum_{k>n}c_k\right|^2\right\} \\ &\le 2\pi\left\{\theta n^{-m}\sum_{1\le k\le\sqrt{n}} g_k^2+\sum_{k>\sqrt{n}} c_k^2\right\}\to 0, n\to\infty. \end{aligned}$$

Note that the linear space invariant map $T_n: f\to T_n(f;x)$ can be evaluated by FFT and is applicable separably in any of the x, y and z-directions. Also note that $\|T_n\|=1$.

We study the effect of this filtering for different θ and m and also the degree of approximation of $f(x)$ by the polynomials $T_n(f;x)$. Let $L_{2\pi,2m}^2$, $(m\ge 1)$ denote the subspace of $L_{2\pi}^2$ consisting of functions $f(x)$ for which

$$\|f\|_{2m}=\left\{\pi\left\{a_0^2/2+\sum_{k\ge 1}k^{2m}\left(a_k^2+b_k^2\right)\right\}\right\}^{1/2}<\infty.$$

With this norm $L_{2\pi,2m}^2$ is a Banach space, and we define the associated Peetre's K-functional by

$$K_{2m}(f;t)=\inf_{g\in L_{2\pi,2m}^2}\left\{\|f-g\|+t^{2m}\|g\|_{2m}\right\}\quad f\in L_{2\pi}^2.$$

Since all trigonometric polynomials belong to $L_{2\pi,2m}^2$, it is dense in $L_{2\pi}^2$, from which it follows that for all $f\in L_{2\pi}^2$, $K_{2m}(f;t)\to 0$, as $t\to 0$.

Lemma 2.2.2.

If $f\in L_{2\pi,2m}^2$, $\|f-T_n(f)\|\le\max(1,\theta)n^{-2m}\|f\|_{2m}$.

Proof: Since,

$$\begin{aligned} f(x)-T_n(f;x) &= \sum_{1\le k\le n}\left(a_k(f)\cos kx+b_k(f)\sin kx\right)\left(\theta(k/n)^{2m}\right)/\left(1+\theta(k/n)^{2m}\right) \\ &\quad+\sum_{k>n}\left(a_k(f)\cos kx+b_k(f)\sin kx\right), \\ \|f-T_n(f)\| &\le \pi\left\{\sum_{1\le k\le n}\left(\|a_k(f)\|^2+\|b_k(f)\|^2\right)\theta(k/n)^{2m}\right. \\ &\quad\left.+\sum_{k>n}\left(\|a_k(f)\|^2+\|b_k(f)\|^2\right)\right\}^{1/2} \\ &\le \max(1,\theta)n^{-2m}\|f\|_{2m}. \end{aligned}$$

The main direct and inverse results on the degree of approximation of $f(x)$ by $T_n(x)$ are summarized in the following:

Theorem 2.2.3.

Let $f \in L_{2\pi}{}^2$, and $\psi \in S$. Then:

(i) There exists a constant $M \leq \max(2,\theta)$, such that

$$f - T_n \leq MK_{2m}(f;1/n) \to 0, n \to \infty;$$

(ii) $|f - T_n = \mathrm{O}(n^{-2m}), \mathrm{n} \to \infty,$ iff $f \in L_{2\pi,2m}{}^2$;

(iii) $\|f - T_n| = \mathrm{o}(n^{-2m}), \mathrm{n} \to \infty,$ iff f is a constant;

(iv) $f - T_n = \mathrm{O}(\psi(n^{-2m})), \mathrm{n} \to \infty,$ iff $K_{2m}(f;t) = \mathrm{O}(\psi(t^{2m})), \mathrm{t} \to 0;$

(v) $f - T_n = \mathrm{o}(\psi(n^{-2m})), \mathrm{n} \to \infty,$ iff $K_{2m}(f;t) = \mathrm{o}(\psi(t^{2m})), \mathrm{t} \to 0;$ and,

(vi) $f - T_n \sim K_{2m}(f;1/n),$ if $K_{2m}(f;t^{1/(2m)}) \in S.$

Proof: Since,

$$f(x) - T_n(x) = \sum_{1\leq k\leq n}\left(a_k \cos kx + b_k \sin kx\right)\left(\theta\, k^{2m}/[n^{2m} + \theta\, k^{2m}]\right) + \sum_{k>n}\left(a_k \cos kx + b_k \sin kx\right),$$

$$f - T_n(f) \leq f - g + |g - T_n(g)| + \|T_n(g - f)| \leq 2\|f - g + \max(1,\theta)n^{-2m}|\, g\|_{2m}.$$

Taking infimum over $g \in L_{2\pi,2m}{}^2$, we have

$$f - T_n(f) \leq M_\theta K_{2m}(f;1/n),$$

where $M_\theta = \max(2,\theta)$. As, $K_{2m}(f;t) \to 0, \mathrm{t} \to 0$, (i) follows.

If $f \in L_{2\pi,2m}{}^2$, by the Lemma, $f - T_n\| = \mathrm{O}(n^{-2m}), \mathrm{n} \to \infty$. Conversely, assuming $f - T_n = \mathrm{O}(n^{-2m})$, if l is a fixed natural number, for all $n \geq l$ sufficiently large, there exists a constant M such that

$$M \geq n^{2m} f - T_n$$

$$\geq n^{2m} \sum_{1\leq k\leq l}\left(a_k(f)\cos kx + b_k(f)\sin kx\right)\theta(k/n)^{2m}\left[1+\theta(k/n)^{2m}\right]^{-1}$$

$$\geq \sum_{1\leq k\leq l}\left(a_k(f)\cos kx + b_k(f)\sin kx\right)(\theta\, k^{2m})\left[1+\theta(k/n)^{2m}\right]^{-1}.$$

Taking limit as $n \to \infty$, we have

$$\sum_{1 \le k \le l} (a_k(f)\cos kx + b_k(f)\sin kx)(\theta\, k^{2m}) \le \mathrm{M}.$$

Letting $l \to \infty$, we conclude $f \in L_{2\pi,2m}{}^2$. This completes the proof of (ii).

For (iii), if f is a constant, $f - T_n = 0 = \mathrm{o}(n^{-2m})$. Conversely, if $f - T_n = \mathrm{o}(n^{-2m})$, as in the proof of (ii), we get

$$\sum_{k \ge 1} (a_k(f)\cos kx + b_k(f)\sin kx)(\theta\, k^{2m}) = 0.$$

Hence, $a_k(f) = b_k(f) = 0$, for all $k \ge 1$, and $f(x)$ must be a constant.

To prove (iv)-(vi), if $K_{2m}(f;t) = \mathrm{O}(\psi(t^{2m})), \mathrm{t} \to 0$, $f - T_n = \mathrm{O}(\psi(n^{-2m})), \mathrm{n} \to \infty$, and if $K_{2m}(f;t) = \mathrm{o}(\psi(t^{2m})), \mathrm{t} \to 0$, $\|f - T_n\| = \mathrm{o}(\psi(n^{-2m})), \mathrm{n} \to \infty$, follows from (i).

For the rest, assuming $f - T_n = \mathrm{O}(\psi(n^{-2m})), \mathrm{n} \to \infty$,

$$\begin{aligned} K_{2m}(f;t) &\le K_{2m}(f - T_n(f);t) + K_{2m}(T_n(f-g);t) + K_{2m}(T_n(g);t) \\ &\le \mathrm{O}(\psi(n^{-2m})) + t^{2m} n^{2m} \|f - g\| + \|g\|_{2m}. \end{aligned}$$

Taking infimum over $g \in L_{2\pi,2m}{}^2$,

$$K_{2m}(f;t) \le \mathrm{O}(\psi(n^{-2m})) + (tn)^{2m}\left(K_{2m}(f;n^{-1})\right),$$

and the results follow from Lemma 1.2.2.

Taking $\psi(t) = t^{\alpha/(2m)}$, for the Bernstein-Lipschitz orders $n^{-\alpha}$, we have:

Corollary 2.2.4.

Let $f \in L_{2\pi}{}^2$, and $0 < \alpha < 2m$. Then:

(i) $f - T_n = \mathrm{O}(n^{-\alpha}), n \to \infty$, iff $K_{2m}(f;t) = \mathrm{O}(t^{\alpha}), t \to 0$;

(ii) $f - T_n = \mathrm{o}(n^{-\alpha}), n \to \infty$, iff $K_{2m}(f;t) = \mathrm{o}(t^{\alpha}), t \to 0$; and,

(iii) $f - T_n \sim n^{-\alpha}$, if $K_{2m}(f;t) \sim t^{\alpha}$.

Note that in the MRI reconstructions using FFT, the relevant choice $\sigma_n = 2^{-n}$ satisfies $\sigma_{n+1} \ge (1/2)\sigma_n$, so that in the converse parts a given degree of

approximation along the sequence $\{2^n\}$ itself implies the corresponding smoothness of f as given in the theorem and the corollary.

2.3. THE BUTTERWORTH FILTER AND GENERALIZATIONS

A lowpass filter produces no or little modification in the Fourier coefficients corresponding to frequencies below a given threshold called the cutoff frequency. It allows the low frequency components pass and reduces or blocks high frequency components. The equations for the transfer functions of the Butterworth low pass filters of 1^{st} and the 2^{nd} order are given by $T(f)=1/(1+(r^2/cf^2))$ and $T(f)=1/(1+(r^4/cf^4))$, respectively, where cf is the cutoff frequency and r is the square root of the sum of the squares of the frequency components. The cutoff cf is normally in the interval [0, 0.5]. The second order filter is steeper than the first order filter.

The minimizing differential filter of Section 2.2 may be termed one-dimensional Butterworth filter of order m. A differential characterization of the two-dimensional Butterworth filters may be made as follows. Let $T_{n,2}$ denote the space of trigonometric polynomials of order n, $n \geq 0$, in the two variables x, y. Let k stand for a double index (k_x, k_y) and $|k|^2 = k_x^2 + k_y^2$. For members of $T_{n,2}$, then, $|k|^2 \leq n^2$. We will consider the space $L_{2\pi,2}{}^2$ of 2π-periodic functions square integrable on the square $[-\pi,\pi]\times[-\pi,\pi]$ with the norm given by

$$\|f\| = \left(\int_{[-\pi,\pi]\cdot[-\pi,\pi]} f(x,y)^2 \, dxdy \right)^{1/2}.$$

Theorem 2.3.1.

For $f \in L_{2\pi,2}{}^2$, $m = 1, 2, 3, \ldots$, and $\theta > 0$, the trigonometric polynomial $T_n(x,y) \in T_{n,2}$ minimizing

$$Q_m = (2\pi)^{-2} \iint_{[-\pi,\pi]\times[-\pi,\pi]} \Big\{ f(x,y) - T_n(x,y)\Big|^2 + \theta \sum_{0\le j\le m} {}^mC_j\, n^{-m} \partial^m T_n(x,y)/\partial x^{m-j} \partial y^j \Big|^2 \Big\} dxdy,$$

is unique and is given by:

$$T_n(x,y) = T_n(f;x,y) = \sum_{1\le|k|\le n} c_k(f) \big[1+\theta(|k|/n)^{2m}\big]^{-1} \exp\{i(k_x x + k_y y)\},$$

where $c_k(f)$ are the complex Fourier coefficients of the function $f(x,y)$. The map: $f \to T_n$ is linear, and, $T_n - f \to 0$, $n \to \infty$.

Proof: Writing

$$f(x,y) = \sum_{k\ge 0} c_k \exp\{i(k_x x + k_y y)\},$$

and

$$T_n(x,y) = \sum_{|k|\le n} g_k \exp\{i(k_x x + k_y y)\},$$

by the orthogonality of the functions $\exp\{i(k_x x + k_y y)\}$ over $[-\pi,\pi]\times[-\pi,\pi]$,

$$\begin{aligned} Q_m &= \sum_{|k|\le n} \Big(|c_k - g_k|^2 + (\theta/n^{2m}) \sum_{0\le j\le m} {}^mC_j (k_x)^{2(m-j)} (k_y)^{2j} |g_k|^2 \Big) + \sum_{|k|>n} |c_k|^2 \\ &= \sum_{|k|\le n} \Big(|c_k - g_k|^2 + (\theta/n^{2m}) |k|^{2m} |g_k|^2 \Big) + \sum_{|k|>n} |c_k|^2. \end{aligned}$$

As $\sum_{|k|>n} |c_k|^2$ does not affect the minimization, each $\Big(|c_k - g_k|^2 + \theta(k/n)^{2m} |g_k|^2 \Big)$ must separately be a minimum. If $c_k = 0$, $g_k = 0$. Else, putting $g_k = \phi c_k$,

$$g(\phi) = 1 - \phi^2 + \theta(|k|/n)^{2m} \phi^2$$

is minimized by a real ϕ satisfying

$$g'(\phi) = 2\Big(\phi\big\{1+\theta(|k|/n)^{2m}\big\} - 1\Big) = 0,$$

and is uniquely given by:

$$\phi = \big[1+\theta(|k|/n)^{2m}\big]^{-1}.$$

Since $g''(\phi) = 2\big(1+\theta(|k|/n)^{2m}\big) > 0$, it is indeed a minima.

It follows that the trigonometric polynomial $T_n(x,y)$ minimizing Q_m is given as:

$$T_n(x,y) = T_n(f;x,y) = \sum_{k\ge 0} c_k \exp\{i(k_x x + k_y y)\} / [1+\theta(|k|/n)^{2m}],$$

with

$$c_k(f) = (2\pi)^{-2} \int_{[-\pi,\pi]\times[-\pi,\pi]} f(x,y)\exp\{i(k_x x + k_y y)\}dxdy,$$

being the Fourier coefficients of $f(x,y)$. For the convergence,

$$\|f - T_n\| = (2\pi)^2 \left\{\sum_{|k|\le n} \theta(|k|/n)^{2m}\left[1+\theta(|k|/n)^{2m}\right]^{-1}|g_k|^2 + \sum_{|k|>n}|c_k|^2\right\}$$
$$\le (2\pi)^2 \left\{\theta n^{-m}\sum_{|k|\le\sqrt{n}}|g_k|^2 + \sum_{|k|>\sqrt{n}}|c_k|^2\right\} \to 0, n\to\infty.$$

The linear space invariant map $T_n : f \to T_n(f;x,y)$ has $\|T_n\| = 1$, and it can be evaluated by FFT using multiplicative replacements: $c_k \to c_k/[1+\theta(|k|/n)^{2m}]$ in the frequency domain. The effect of this filtering for different θ and m and also the degree of approximation of $f(x,y)$ by the polynomials $T_n(f;x,y)$ is considered below.

Let $L_{2\pi,2,2m}{}^2$, (m ≥ 1) denote the subspace of $L_{2\pi,2}{}^2$ of functions $f(x,y)$ for which

$$|f\|_{2m} = \left(4\pi^2\left\{|c_o|^2 + \sum_{|k|\ge 1}|k|^{2m}|c_k|^2\right\}\right)^{1/2} < \infty.$$

With this norm $\|.\|_{2m}$, $L_{2\pi,2,2m}{}^2$ is a Banach space, and the associated Peetre's K-functional for the pair $(L_{2\pi,2}{}^2, L_{2\pi,2,2m}{}^2)$ of spaces is given by:

$$K_{2m,2}(f;t) = \inf\left\{\|f-g\| + t^{2m}\|g\|_{2m} : g \in L_{2\pi,2,2m}{}^2\right\},\ f \in L_{2\pi,2}{}^2.$$

Since all trigonometric polynomials in two variables belong to $L_{2\pi,2,2m}{}^2$, it is dense in $L_{2\pi,2}{}^2$, and it follows that for all $f \in L_{2\pi,2}{}^2$, $K_{2m,2}(f;t) \to 0$, as $t \to 0$.

Lemma 2.3.2.

If $f \in L_{2\pi,2,2m}{}^2$, $\|f - T_n(f)\| \le \max(1,\theta)n^{-2m}\|f\|_{2m}$.

Proof: We have,

$$f(x,y) - T_n(f;x,y) = \sum_{1\le|k|\le n}\left(c_k(f)\exp\{i(k_x x + k_y y)\}\right)$$
$$\times\left(\theta(|k|/n)^{2m}\right)/\left(1+\theta(|k|/n)^{2m}\right) + \sum_{k>n}\left(c_k(f)\exp\{i(k_x x + k_y y)\}\right).$$

Hence,

$$f - T_n(f) \le 4\pi^2 \left\{ \sum_{1 \le k \le n} \left(c_k(f)^2 \right) \theta(|k|/n)^{2m} \right\}^2 + \sum_{k>n} \left(|c_k(f)^2| \right)^{1/2}$$
$$\le \max(1,\theta) n^{-2m} \| f \|_{2m}.$$

The main direct and inverse results on the degree of approximation of $f(x,y)$ by $T_n(f;x,y)$ are:

Theorem 2.3.3.

Let $f \in L_{2\pi,2}{}^2$, and $\psi \in S$. Then:

(i) There exists a constant $M \le \max(2,\theta)$, such that

$$f - T_n \le M K_{2m,2}(f;1/n) \to 0, \; n \to \infty;$$

(ii) $f - T_n = O(n^{-2m}), n \to \infty$, iff $f \in L_{2\pi,2,2m}{}^2$;

(iii) $f - T_n = o(n^{-2m}), n \to \infty$, iff f is a constant;

(iv) $f - T_n = O(\psi(n^{-2m})), n \to \infty$, iff $K_{2m,2}(f;t) = O(\psi(t^{2m})), t \to 0$;

(v) $f - T_n = o(\psi(n^{-2m})), n \to \infty$, iff $K_{2m,2}(f;t) = o(\psi(t^{2m})), t \to 0$; and,

(vi) $f - T_n \sim K_{2m,2}(f;1/n)$, if $K_{2m,2}(f;t^{1/(2m)}) \in S$.

Proof: Since,

$$f(x,y) - T_n(f;x,y) = \sum_{1 \le |k| \le n} \left(c_k(f) \exp\{i(k_x x + k_y y)\} \right)$$
$$\times \left(\theta(|k|/n)^{2m} \right) / \left(1 + \theta(|k|/n)^{2m} \right) + \sum_{k>n} \left(c_k(f) \exp\{i(k_x x + k_y y)\} \right),$$

$$f - T_n(f) \le \| f - g \| + \| g - T_n(g) \| + \| T_n(g - f) \|$$
$$\le 2 \| f - g \| + \max(1,\theta) n^{-2m} \| g \|_{2m}.$$

Taking infimum over $g \in L_{2\pi,2,2m}{}^2$, we have

$$f - T_n(f) \le M_\theta K_{2m,2}(f;1/n),$$

where $M_\theta = \max(2,\theta)$. As, $K_{2m,2}(f;t) \to 0, t \to 0$, (i) follows. If $f \in L_{2\pi,2,2m}{}^2$, by the previous Lemma, $\| f - T_n \| = O(n^{-2m}), n \to \infty$.

Conversely, assuming $\|f-T_n\| = O(n^{-2m})$, if l is a fixed natural number, for all $n \geq l$ sufficiently large, there exists a constant M such that

$$\begin{aligned} M &\geq n^{2m}\|f-T_n\| \\ &\geq n^{2m}\sum_{1\leq k\leq l}\left(c_k(f)\exp\{i(k_x x+k_y y)\}\right)\theta(|k|/n)^{2m}\left[1+\theta(|k|/n)^{2m}\right]^{-1} \\ &\geq \sum_{1\leq k\leq l}\left(c_k(f)\exp\{i(k_x x+k_y y)\}\right)(\theta|k|^{2m})\left[1+\theta(k/n)^{2m}\right]^{-1}. \end{aligned}$$

Taking limit as $n\to\infty$, we have

$$\sum_{1\leq k\leq l}\left(c_k(f)\exp\{i(k_x x+k_y y)\}\right)(\theta k^{2m}) \leq M.$$

Letting $l\to\infty$, we conclude that $f\in L_{2\pi,2,2m}{}^2$. This completes the proof of (ii).

For (iii), if f is a constant, $\|f-T_n\| = 0 = o(n^{-2m})$. Conversely, if $\|f-T_n\| = o(n^{-2m})$, as in the proof of (ii), we get

$$\sum_{k\geq 1}\left(c_k(f)\exp\{i(k_x x+k_y y)\}\right)(\theta k^{2m}) = 0.$$

Hence, $c_k(f)=0$, for all $k\geq 1$, and f must be a constant.

To prove (iv)-(vi), if $K_{2m,2}(f;t) = O(\psi(t^{2m})), t\to 0$, $|f-T_n| = O(\psi(n^{-2m})), n\to\infty$, and if $K_{2m,2}(f;t) = o(\psi(t^{2m})), t\to 0$, $\|f-T_n\| = o(\psi(n^{-2m})), n\to\infty$, follows from (i). For the rest, assuming $\|f-T_n\| = O(\psi(n^{-2m})), n\to\infty$,

$$\begin{aligned} K_{2m,2}(f;t) &\leq K_{2m,2}(f-T_n(f);t)+K_{2m,2}(T_n(f-g);t)+K_{2m,2}(T_n(g);t) \\ &\leq O(\psi(n^{-2m}))+t^{2m}n^{2m}\|f-g\|_{2m}. \end{aligned}$$

Taking infimum over $g\in L_{2\pi,2,2m}{}^2$,

$$K_{2m,2}(f;t) \leq O(\psi(n^{-2m}))+(tn)^{2m}\left(K_{2m,2}(f;n^{-1})\right),$$

and the results follow from Lemma 1.2.2.

Taking $\psi(t)=t^{\alpha/(2m)}$, for the Bernstein-Lipschitz orders $n^{-\alpha}$, we have:

Corollary 2.3.4.

Let $f\in L_{2\pi,2}{}^2$, and $0<\alpha<2m$. Then:

(i) $\|f-T_n\| = O(n^{-\alpha}), n\to\infty$, iff $K_{2m,2}(f;t) = O(t^{\alpha}), t\to 0$;

(ii) $f - T_n = o(n^{-\alpha}), n \to \infty$, iff $K_{2m,2}(f;t) = o(t^{\alpha}), t \to 0$; and,

(iii) $f - T_n \sim n^{-\alpha}$, if $K_{2m,2}(f;t) \sim t^{\alpha}$.

In the MRI reconstructions using FFT, the relevant choice $\sigma_n = 2^{-n}$ satisfies $\sigma_{n+1} \geq (1/2)\sigma_n$, so that in the converse parts a given degree of approximation along the sequence $\{2^n\}$ itself implies the corresponding smoothness of f as given in the theorem and the corollary.

2.4. THE 3-D BUTTERWORTH FILTER

Following the differential characterization of the 2-D Butterworth filters, a differential characterization of 3-D Butterworth filters is considered in this section. Let $T_{n,3}$ denote the space of trigonometric polynomials of order n, $n \geq 0$, in the three variables x, y, z. Let k stand for a triple index (k_x, k_y, k_z) and $|k|^2 = k_x^2 + k_y^2 + k_z^2$. For members of $T_{n,3}$, then, $|k|^2 \leq n^2$. We consider the space $L_{2\pi,3}^2$ of 2π-periodic functions square integrable on the cube $[-\pi,\pi]\times[-\pi,\pi]\times[-\pi,\pi]$ normed by

$$\|f\| = \left(\int_{[-\pi,\pi]\cdot[-\pi,\pi]\cdot[-\pi,\pi]} f(x,y,z)^2 \, dxdydz \right)^{1/2}.$$

Theorem 2.4.1.

For $f \in L_{2\pi,3}^2$, $m = 1, 2, 3, \ldots$, and $\theta > 0$, the trigonometric polynomial $T_n(x,y,z) \in T_{n,3}$ minimizing

$$Q_m = (2\pi)^{-3} \int_{[-\pi,\pi]\cdot[-\pi,\pi]\cdot[-\pi,\pi]} \Big\{ f(x,y,z) - T_n(x,y,z)^2 + \theta \sum_{0 \leq i,j,l; i+j+l=m} [m!/(i!\,j!\,l!)] n^{-m} \partial^m T_n(x,y,z)/\partial x^i \partial y^j \partial z^l \Big\}^2 dxdydz,$$

is unique and is given by:

$$T_n(x,y,z) = T_n(f;x,y,z) = \sum_{1 \leq |k| \leq n} c_k(f) \left[1 + \theta(|k|/n)^{2m}\right]^{-1} \exp\{i(k_x x + k_y y + k_z z)\},$$

where $c_k(f)$ are the complex Fourier coefficients of the function $f(x,y,z)$. The map: $f \to T_n$ is linear, and, $T_n - f \to 0$, $n \to \infty$.

Proof: Writing

$$f(x,y,z) = \sum_{k \geq n} c_k \exp\{i(k_x x + k_y y + k_z z)\},$$

and

$$T_n(x,y,z) = \sum_{|k| \leq n} g_k \exp\{i(k_x x + k_y y + k_z z)\}.$$

by the orthogonality of the functions $\exp\{i(k_x x + k_y y + k_z z)\}$ over $[-\pi,\pi]\times[-\pi,\pi]\times[-\pi,\pi]$,

$$Q_m = \sum_{|k| \leq n} \left(c_k - g_k \right)^2 + (\theta / n^{2m}) \sum_{0 \leq i,j,l; i+j+l=m} [m!/(i!\, j!\, l!)] (k_x)^{2i} (k_y)^{2j} (k_z)^{2l} \left| g_k \right|^2 \Big) + \sum_{|k|>n} \left| c_k \right|^2$$

$$= \sum_{|k| \leq n} \left(c_k - g_k \right)^2 + (\theta / n^{2m}) \left| k \right|^{2m} \left| g_k \right|^2 \Big) + \sum_{|k|>n} \left| c_k \right|^2.$$

To minimize Q_m, each $\left(c_k - g_k \right)^2 + \theta (k/n)^{2m} \left| g_k \right|^2 \Big)$ must be a minimum. As $g_k = 0$, if $c_k = 0$, with $g_k = \phi\, c_k$,

$$g(\phi) = 1 - \phi^2 + \theta (|k|/n)^{2m} \phi^2$$

is minimized by the real ϕ satisfying

$$g'(\phi) = 2\Big(\phi\{1 + \theta(|k|/n)^{2m}\} - 1\Big) = 0,$$

and is uniquely given by:

$$\phi = 1 + \theta(|k|/n)^{2m}{}^{-1}.$$

Since $g''(\phi) = 2\Big(1 + \theta(|k|/n)^{2m}\Big) > 0$, it is indeed a minima.

Hence the trigonometric polynomial $T_n(x,y,z)$ minimizing Q_m is given as:

$$T_n(x,y,z) = T_n(f;x,y,z) = \sum_{k \geq 0} c_k \exp\{i(k_x x + k_y y + k_z z)\} / [1 + \theta(|k|/n)^{2m}],$$

with

$$c_k(f) = (2\pi)^{-3} \int_{[-\pi,\pi]\times[-\pi,\pi]\times[-\pi,\pi]} f(x,y,z) \exp\{i(k_x x + k_y y + k_z z)\} dx dy dz,$$

being the Fourier coefficients of $f(x,y,z)$. For the convergence,

$$f - T_n = (2\pi)^3 \left(\sum_{|k|\le n} \theta(|k|/n)^{2m} \left[1+\theta(|k|/n)^{2m}\right]^{-1} |g_k|^2 + \sum_{k>n} |c_k|^2 \right)$$
$$= (2\pi)^3 \left\{ \theta\, n^{-m} \sum_{|k|\le\sqrt{n}} |g_k|^2 + \sum_{|k|>\sqrt{n}} |c_k|^2 \right\} \to 0, n \to \infty.$$

As in the 2-D case, the linear space invariant map $T_n : f \to T_n(f;x,y,z)$ has $\|T_n\| = 1$, and can be evaluated by FFT using multiplicative replacements: $c_k \to c_k / \left(1+\theta(|k|/n)^{2m}\right)$ in the frequency domain. The approximation properties of the polynomials $T_n(f;x,y,z)$ are considered below.

Let $L_{2\pi,3,2m}{}^2$, $(m \ge 1)$ denote the subspace of $L_{2\pi,3}{}^2$ of functions $f(x,y,z)$ for which

$$\|f\|_{2m} = \left(8\pi^3 \left\{ |c_o|^2 + \sum_{|k|\ge 1} |k|^{2m} |c_k|^2 \right\}\right)^{1/2} < \infty.$$

With this norm $\|.\|_{2m}$, $L_{2\pi,3,2m}{}^2$ is a Banach space, and the associated Peetre's K-functional for the pair $(L_{2\pi,3}{}^2, L_{2\pi,3,2m}{}^2)$ of spaces is given by:

$$K_{2m,3}(f;t) = \inf \left\{ \|f-g\| + t^{2m} \|g\|_{2m} : g \in L_{2\pi,3,2m}{}^2 \right\} f \in L_{2\pi,3}{}^2.$$

Since all trigonometric polynomials in two variables belong to $L_{2\pi,3,2m}{}^2$, it is dense in $L_{2\pi,3}{}^2$, and it follows that for all $f \in L_{2\pi,3}{}^2$, $K_{2m,3}(f;t) \to 0$, as $t \to 0$.

Lemma 2.4.2.

If $f \in L_{2\pi,3,2m}{}^2$, $\|f - T_n(f)\| \le \max(1,\theta) n^{-2m} \|f\|_{2m}$.

Proof: We have,

$$f(x,y,z) - T_n(f;x,y,z)$$
$$= \sum_{1\le|k|\le n} \left(c_k(f) \exp\{i(k_x x + k_y y + k_z z)\}\right) \left(\theta(|k|/n)^{2m}\right) / \left(1+\theta(|k|/n)^{2m}\right)$$
$$+ \sum_{k>n} \left(c_k(f) \exp\{i(k_x x + k_y y + k_z z)\}\right).$$

Hence,

$$f - T_n(f) \le 8\pi^3 \left\{ \sum_{1 \le k \le n} \left(c_k(f)^2 \right) \theta(|k|/n)^{2m} \right\}^2 + \sum_{k>n} \left(c_k(f)^2 \right) \Big\}^{1/2}$$
$$\le \max(1,\theta) n^{-2m} |f|_{2m}.$$

The main direct and inverse results on the degree of approximation of $f(x,y,z)$ by $T_n(x,y,z)$ are:

Theorem 2.4.3.

Let $f \in L_{2\pi,3}{}^2$, and $\psi \in S$. Then:

(i) There exists a constant $M \le \max(2,\theta)$, such that

$$f - T_n \le MK_{2m,3}(f;1/n) \to 0, n \to \infty;$$

(ii) $f - T_n = O(n^{-2m}), n \to \infty$, iff $f \in L_{2\pi,3,2m}{}^2$;

(iii) $f - T_n = o(n^{-2m}), n \to \infty$, iff f is a constant;

(iv) $f - T_n = O(\psi(n^{-2m})), n \to \infty$, iff $K_{2m,3}(f;t) = O(\psi(t^{2m})), t \to 0$;

(v) $f - T_n = o(\psi(n^{-2m})), n \to \infty$, iff $K_{2m,3}(f;t) = o(\psi(t^{2m})), t \to 0$; and,

(vi) $f - T_n \sim K_{2m,3}(f;1/n)$, if $K_{2m,3}(f;t^{1/(2m)}) \in S$.

Proof: As:

$$f(x,y,z) - T_n(f;x,y)$$
$$= \sum_{1 \le |k| \le n} \left(c_k(f) \exp\{i(k_x x + k_y y + k_z z)\} \right) \left(\theta(|k|/n)^{2m} \right) / \left(1 + \theta(|k|/n)^{2m} \right)$$
$$+ \sum_{k>n} \left(c_k(f) \exp\{i(k_x x + k_y y + k_z z)\} \right),$$

$$f - T_n(f) \le f - g + g - T_n(g) + T_n(g - f)$$
$$\le 2|f - g| + \max(1,\theta) n^{-2m} |g|_{2m}.$$

Taking infimum over $g \in L_{2\pi,3,2m}{}^2$, we have

$$f - T_n(f) \le M_\theta K_{2m,3}(f;1/n),$$

where $M_\theta = \max(2,\theta)$. As, $K_{2m,3}(f;t) \to 0, t \to 0$, (i) follows.

If $f \in L_{2\pi,3,2m}{}^2$, by the previous Lemma, $|f - T_n| = O(n^{-2m}), n \to \infty$.

Conversely, assuming $\|f-T_n\| = O(n^{-2m})$, if l is a fixed natural number, for all $n \geq l$ sufficiently large, there exists a constant M such that

$$\begin{aligned} M &\geq n^{2m}\|f-T_n\| \\ &\geq n^{2m}\|\sum_{1\leq k\leq l}\left(c_k(f)\exp\{i(k_x x+k_y y+k_z z)\}\right)\theta(|k|/n)^{2m}\,1+\theta(|k|/n)^{2m}\|^{-1} \\ &\geq \|\sum_{1\leq k\leq l}\left(c_k(f)\exp\{i(k_x x+k_y y+k_z z))\}\right)(\theta|k|^{2m})\,1+\theta(k/n)^{2m}\|^{-1}. \end{aligned}$$

Taking limit as $n\to\infty$, we have

$$\|\sum_{1\leq k\leq l}\left(c_k(f)\exp\{i(k_x x+k_y y+k_z z))\}\right)(\theta k^{2m})\| \leq M.$$

Letting $l\to\infty$, we conclude that $f\in L_{2\pi,3,2m}{}^2$. This completes the proof of (ii).

For (iii), if f is a constant, $\|f-T_n\| = 0 = o(n^{-2m})$. Conversely, if $\|f-T_n\| = o(n^{-2m})$, as in the proof of (ii), we get

$$\|\sum_{k\geq 1}\left(c_k(f)\exp\{i(k_x x+k_y y+k_z z)\}\right)(\theta k^{2m})\| = 0.$$

Hence, $c_k(f)=0$, for all $k\geq 1$, and f must be a constant.

To prove (iv)-(vi), if $K_{2m,3}(f;t) = O(\psi(t^{2m})), t\to 0$, $\|f-T_n\| = O(\psi(n^{-2m})), n\to\infty$, and if $K_{2m,3}(f;t) = o(\psi(t^{2m})), t\to 0$, $\|f-T_n\| = o(\psi(n^{-2m})), n\to\infty$, follows from (i). For the rest, assuming $\|f-T_n\| = O(\psi(n^{-2m})), n\to\infty$,

$$\begin{aligned} K_{2m,3}(f;t) &\leq K_{2m,3}(f-T_n(f);t)+K_{2m,3}(T_n(f-g);t)+K_{2m,3}(T_n(g);t) \\ &\leq O(\psi(n^{-2m}))+t^{2m}n^{2m}\|f-g\|+\|g\|_{2m}. \end{aligned}$$

Taking infimum over $g\in L_{2\pi,3,2m}{}^2$,

$$K_{2m,3}(f;t) \leq O(\psi(n^{-2m}))+(tn)^{2m}\left(K_{2m,3}(f;n^{-1})\right),$$

and the results follow from Lemma 1.2.2.

Taking $\psi(t)=t^{\alpha/(2m)}$, for the Bernstein-Lipschitz orders $n^{-\alpha}$, we have:

Corollary 2.4.4.

Let $f\in L_{2\pi,3}{}^2$, and $0<\alpha<2m$. Then:

(i) $\|f-T_n\| = O(n^{-\alpha}), n\to\infty$, iff $K_{2m,3}(f;t) = O(t^{\alpha}), t\to 0$;

(ii) $|f - T_n = o(n^{-\alpha}), n \to \infty$, iff $K_{2m,3}(f;t) = o(t^{\alpha}), t \to 0$; and,

(iii) $f - T_n \sim n^{-\alpha}$, if $K_{2m,3}(f;t) \sim t^{\alpha}$.

In the MRI reconstructions using FFT, the relevant choice $\sigma_n = 2^{-n}$ satisfies $\sigma_{n+1} \geq (1/2)\sigma_n$, so that in the converse parts a given degree of approximation along the sequence $\{2^n\}$ itself implies the corresponding smoothness of f as given in the theorem and the corollary.

2.5. NUMERICAL SIMULATIONS AND DISCUSSION

In this section, numerical simulation results of approximation by two-dimensional Butterworth filter are presented. All the experiments in this section are done on test image set II.

Figures 2.1-2.3 show segmented images obtained by the application of the multi-resolution MLE-based segmented images on PD, T_2 and T_1 weighted images approximated with two-dimensional Butterworth filter with $\theta = 1, 2, 4, 8, 16, 32$ for the values of $m = 1, 2$ and 3. The number of iterations taken with $\theta = 1, 2, 4, 8, 16, 32$ for $m = 1$ are 57, 86, 59, 32, 45 and 70 respectively. The number of iterations taken with $\theta = 1, 2, 4, 8, 16, 32$ for $m = 2$ are 73, 49, 50, 53, 57 and 72 respectively. The number of iterations taken with $\theta = 1, 2, 4, 8, 16, 32$ for $m = 3$ are 53, 53, 53, 54, 49 and 50 respectively.

θ	m=1	m=2	m=3
1	0.723859	0.091625	0.015913
2	1.314130	0.177635	0.031546
4	2.249419	0.335402	0.062014
8	3.591323	0.606762	0.120002
16	5.362906	1.033300	0.225807
32	7.584896	1.637128	0.406243

Table 2.1: % Relative L1 Errors in the Approximation of Original PD weighted Image with 2-D Butterworth Filter

θ	m=1	m=2	m=3
1	0.968174	0.142415	0.026660
2	1.730013	0.274785	0.052799
4	2.890335	0.514285	0.103590
8	4.462893	0.916510	0.199689
16	6.385664	1.525403	0.373099
32	8.597831	2.343898	0.662873

Table 2.2: % Relative L1 Errors in the Approximation of Original T_2 weighted Image with 2-D Butterworth Filter

θ	m=1	m=2	m=3
1	0.719274	0.091599	0.016963
2	1.309699	0.176995	0.033582
4	2.256373	0.332299	0.065835
8	3.645170	0.596168	0.126734
16	5.534366	1.005681	0.236241
32	7.991690	1.582371	0.418477

Table 2.3: % Relative L1 Errors in the Approximation of Original T_1 weighted Image with 2-D Butterworth Filter

θ	m=1	m=2	m=3
1	0.841595	0.089839	0.014514
2	1.558778	0.175110	0.028807
4	2.749586	0.333892	0.056763
8	4.562928	0.614222	0.110335
16	7.095091	1.073159	0.209350
32	10.370586	1.759173	0.382140

Table 2.4: % Relative L2 Errors in the Approximation of Original PD weighted Image with 2-D Butterworth Filter

θ	m=1	m=2	m=3
1	0.897747	0.116682	0.021314
2	1.634145	0.225710	0.042227
4	2.810198	0.424509	0.082908
8	4.512003	0.763304	0.160048
16	6.748191	1.289840	0.299839
32	9.461606	2.029814	0.535309

Table 2.5: % Relative L2 Errors in the Approximation of Original T_2 weighted Image with 2-D Butterworth Filter

θ	m=1	m=2	m=3
1	0.901670	0.089823	0.015247
2	1.681128	0.174589	0.030208
4	2.997882	0.331448	0.059309
8	5.050265	0.606569	0.114506
16	7.985939	1.057225	0.214654
32	11.858933	1.744758	0.384177

Table 2.6: % Relative L2 Errors in the Approximation of Original T_1 weighted Image with 2-D Butterworth Filter

The percentage relative errors in the approximations of the two-dimensional Butterworth filter applied to PD, T_2 and T_1 weighted images with $\theta = 1, 2, 4, 8, 16, 32$ corresponding to the values of $m = 1, 2$ and 3 are given in Tables 2.1 - 2.6.

Graph 2.1 (% Relative L1 errors in Approximation with two-dimensional Butterworth filter (m=1,2,3)) and Graph 2.2 (% Relative L2 errors in Approximation with two-dimensional Butterworth filter (m=1,2,3)) present the behavior of % relative errors with respect to the values of m in each of the weighted images. It is clear from the plots that smoothing increases with the increase in the value of θ and decreases with the increase in the value of m as mentioned in section 2.1.

In all the segmented images, the total number of segments obtained are thirteen and visually they are indistinguishable from each other. Out of all these segmented images, one would select that particular image which gives the better contrast between the segments and the overall sizes of each and every segment matches with that represented by given original PD, T_2 and T_1 weighted images. White matter and gray matter are well separated by a boundary in each of these images.

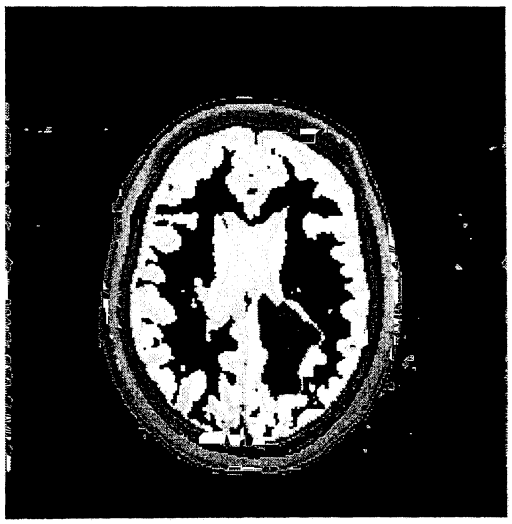
θ=1

θ=2

θ=4

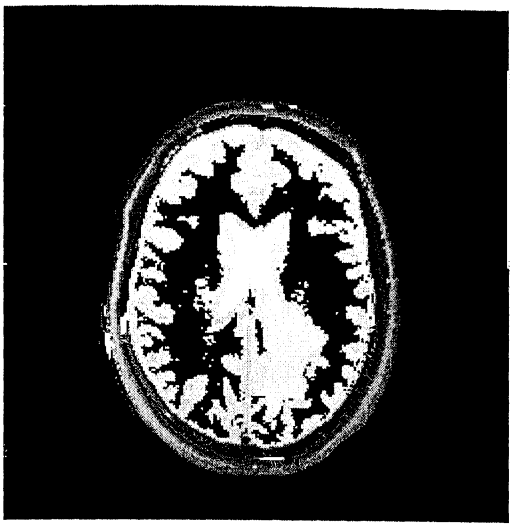
θ=8

θ=16



θ=32

Figure 2.1: MLE Based Segmentation of Images Approximated with 2-D Butterworth Filter (m=1)



θ=1

θ=2

θ=4

θ=8

θ=16

θ=32

Figure 2.2: MLE Based Segmentation of Images Approximated with 2-D Butterworth Filter (m=2)

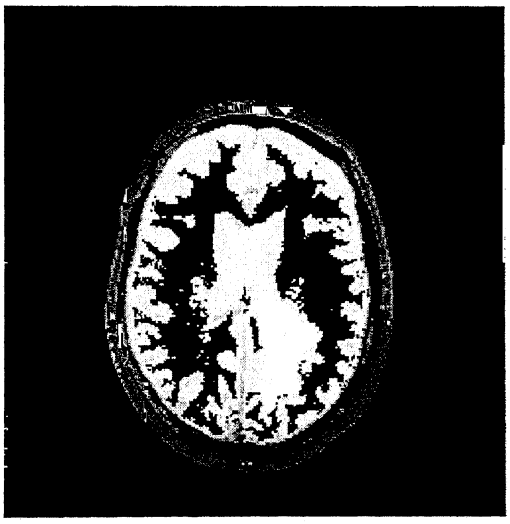
θ=1

θ=2

θ=4

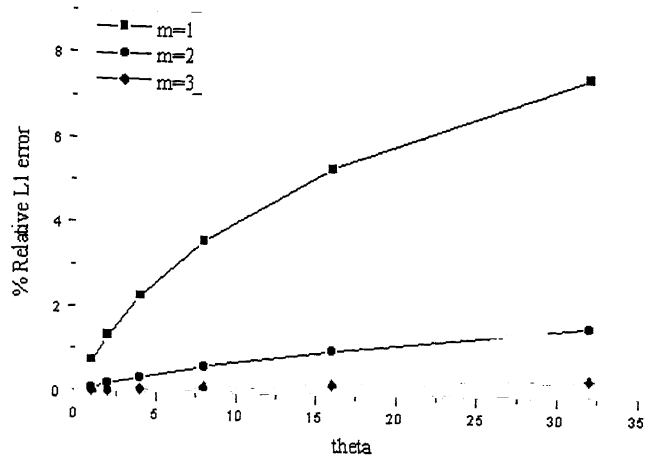
θ=8

θ=16

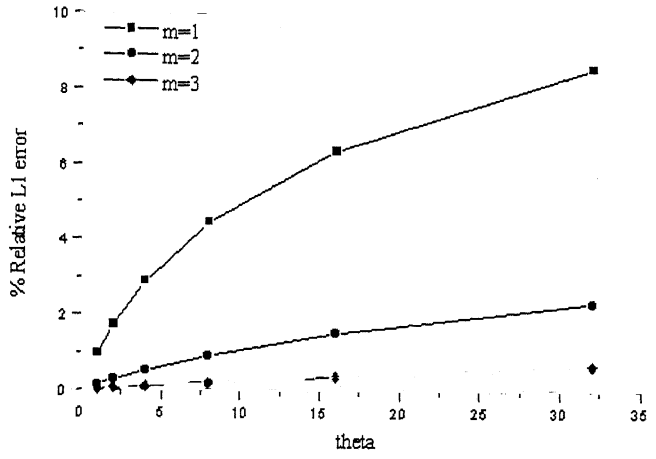


θ=32

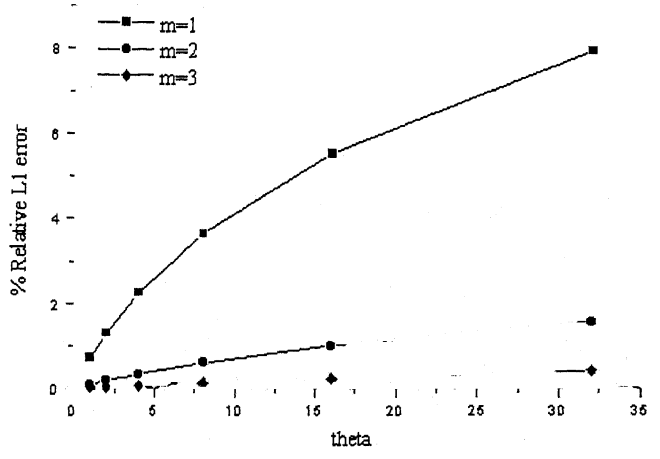
Figure 2.3: MLE Based Segmentation of Images Approximated with 2-D Butterworth Filter (m=3)



Proton Density weighted Image

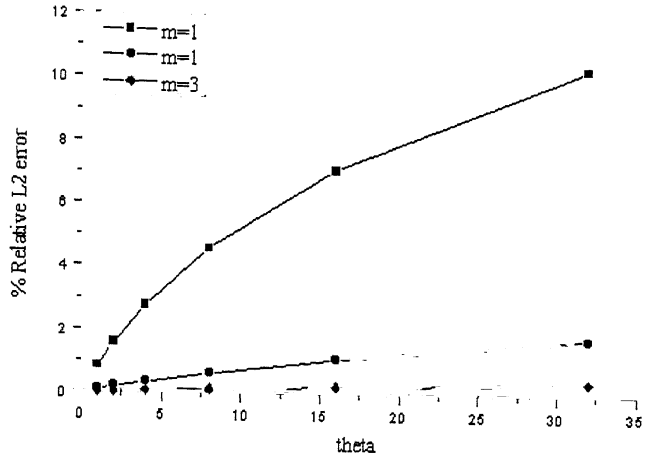


T_2 weighted Image

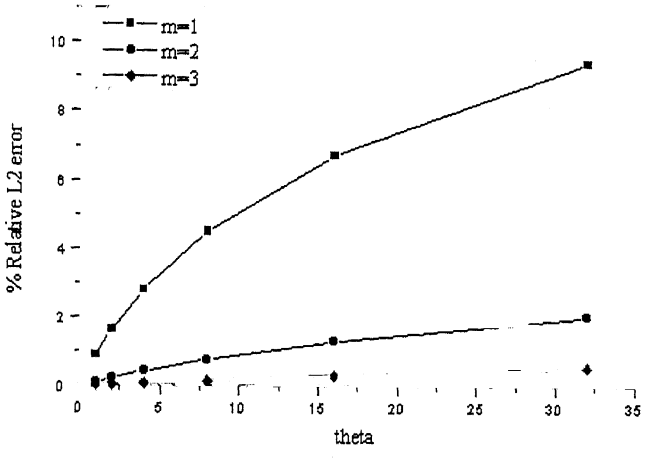


T_1 weighted Image

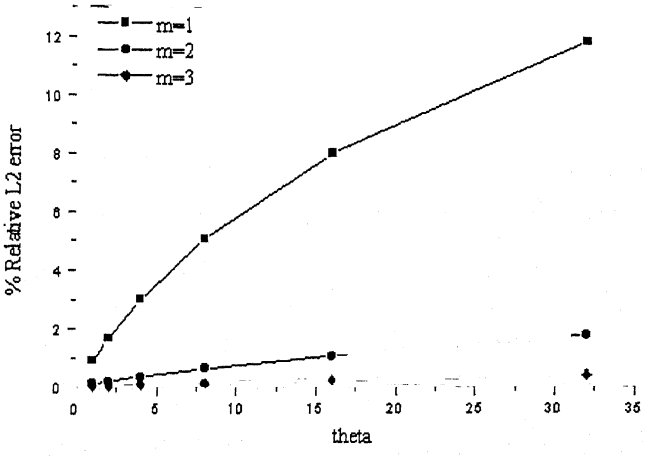
Graph 2.1: % Relative L1 Errors in Approximation with 2-D Butterworth Filter (m=1,2,3)



Proton Density weighted Image



T_2 weighted Image



T_1 weighted Image

Graph 2.2: % Relative L2 Errors in Approximation with 2-D Butterworth Filter (m=1,2,3)

CHAPTER III : DISCRETE CONVOLUTIONS FOR SEGMENTATION

This chapter is a study of approximations and the resulting segmentations by discrete convolutions with Korovkin, Jackson, generalized Jackson and De La Vallée-Poussin filters. Sections 3.1-2 contain the basic results required for the subsequent analysis. Section 3.3 contains the direct and inverse results on the approximation of Korovkin, Jackson and generalized Jackson operators. Discrete convolutions with the De La Vallée-Poussin kernels have been studied in Section 3.4.

Numerical results corresponding to Korovkin, Jackson and generalized Jackson operators, along with their comparisons, have been considered in Section 3.5. Section 3.6 includes numerical simulations with De La Vallée-Poussin kernels.

3.1. INTRODUCTION

Convolutions such as with a Gaussian filter have discretization errors in their implementations. This necessitates a further error analysis and care during their implementations. Also, the Gaussian filters $G(X,\sigma)$ do a spherically symmetric isotropic filtering.

However, the rectangular image size, as well as the nature of noise in the phase and the frequency directions in MRI may necessitate applying cartesian products $L_l(L_m(L_n(f;x);y);z)$ (3-D), $L_m(L_n(f;x);y)$ (2-D), and, $L_n(f;x)$ (1-D) of possibly different filters in the x, y, and z directions.

The degree of approximation results in such separable cases are essentially one dimensional. For example, in 3-D with

$$\|f\| = \max\{|f(x,y,z)| : -\pi \le x,y,z \le \pi\},$$

we have:

Lemma 3.1.1.

If the convolution filters $L_l f$, $M_m f$, $N_n f \to f$, for $f \in C_{2\pi}$,

$$L_l M_m N_n(f) - f = O(1/\varphi_1(l)) + O(1/\varphi_2(m)) + O(1/\varphi_3(n)), (l,m,n \to \infty),$$

iff

$$\|L_l(f) - f\| = O(1/\varphi_1(l)), (l \to \infty),$$

$$M_m(f) - f = O(1/\varphi_2(m)), (m \to \infty),$$

$$N_n(f) - f = O(1/\varphi_3(n)), (n \to \infty).$$

Proof: Since the approximation property for $C_{2\pi}$ implies uniform boundedness, let the constants L, M, N be such that $L_l \le L$, $M_m \le M$, $N_n \le N$. Then, the 1-D results imply

$$\begin{aligned} L_l(M_m(N_n(f;x);y);z) - f(x,y,z) &= L_l(M_m(N_n(f(\xi,\eta,\zeta);x);y);z) - f(x,y,z) \\ &\le L_l(M_m(N_n(f(\xi,\eta,\zeta);x);y);z) - f(((x,\eta,\zeta);y);z) \\ &\quad + L_l(M_m(f(x,\eta,\zeta);y);z) - f((x,y,\zeta);z) \\ &\quad + \|L_l(f(x,y,\zeta);z) - f(x,y,z)\| \\ &= LMO(1/\varphi_3(n)) + MO(1/\varphi_2(m)) + O(1/\varphi_1(l)) \\ &= O(1/\varphi_1(l)) + O(1/\varphi_2(m)) + O(1/\varphi_3(n)). \end{aligned}$$

Conversely, the 1-D results follow from the 3-D relation by letting $m, n \to \infty$, $l, n \to \infty$, and, $l, m \to \infty$.

In view of this Lemma, we shall consider a class of discrete 1-D convolutions associated with continuous convolutions with trigonometric polynomials that can be easily implemented both in the spatial and frequency domains in any of one, two, or three dimensions and observe that they possess essentially the same approximation and filtering properties as the parent continuous trigonometric versions. Moreover, their implementation in the case of two and three variables can be done separably.

Some continuous trigonometric convolution operators are:

De La Vallée-Poussin:

$$V_n(f;x) = [(2n)!!/(2\pi(2n-1)!!)]_{*[-\pi,\pi]}^{*} f(t)\cos^{2n}\left(\frac{t-x}{2}\right)dt.$$

Korovkin: *Fejér-Korovkin*:

$$A_{n-2,l(n)}(f;x) = \int_{[-\pi,\pi]} f(t)\ (1/(\pi n))\left\{\frac{\sin(\pi/n)\cos[n(t-x)/2]}{\cos(t-x)-\cos(\pi/n)}\right\}^2 dt.$$

Jackson:

$$L_{2n-2,l(n)}(f;x) = \int_{[-\pi,\pi]} f(t)\ 3/(2\pi n(2n^2+1))\left\{\frac{\sin[n(t-x)/2]}{\sin[(t-x)/2]}\right\}^4 dt.$$

Generalized Jackson:

$$L_{np-p,l(n)}(f;x) = A^{-1}_{np-p} \int_{[-\pi,\pi]} f(t)\left\{\frac{\sin[n(t-x)/2]}{\sin[(t-x)/2]}\right\}^{2p} dt, (p \geq 3),$$

where $A_{np-p} = \int_{[-\pi,\pi]} [\sin(nt/2)/\sin(t/2)]^{2p} dt.$

For these there hold the asymptotic formulae (see, e.g., P.P. Korovkin [1960], I.P. Natanson [1964], R.A. DeVore [1972], R.K.S. Rathore [1974a]):

$$V_n(f;x) - f(x) = f''(x)/n + o(1/n), n \to \infty, (l(n) > n+2),$$

$$A_{n-2}(f;x) - f(x) = \pi^2 f''(x)/(2n^2) + o(1/n^2), n \to \infty, (l(n) > n),$$

$$L_{2n-2}(f;x) - f(x) = 3f''(x)/(2n^2) + o(1/n^2), n \to \infty, (l(n) > 2n),$$

$$L_{np-p}(f;x) - f(x) = \left(1 - \rho_1^{(np-p)}/\rho_0^{(np-p)}\right) f''(x) + o(1/n^2), n \to \infty, (l(n) > np-p+2),$$

where $[\sin(nt/2)/\sin(t/2)]^{2p} = (1/2)\rho_0^{(np-p)} + \sum_{1 \leq k \leq np-p} \rho_0^{(np-p)} \cos kt.$

Another class of useful convolution filters is given by the Poission operators:

$$L_r(f;x) = [1/(2\pi)] \int_{[-\pi,\pi]} f(t) \left(\frac{(1-r^2)}{(1-2r\cos(t-x)+r^2)}\right) dt,$$

which are positive for $0 < r < 1$.

3.2. DISCRETE CONVOLUTIONS

For purposes of segmentation it is desirable that the filtering does not lead to extraneous, or artifactual segments. This suggests the use of linear positive filters

which by definition do not introduce a zero crossing if the function on which they are applied are non-negative. For convolutions:

$$L_n(f;x) = (1/\pi)\dot{\int}_{[-\pi,\pi]} f(t)K_n(t-x)dt,$$

with general even trigonometric kernels of a form

$$K_n(x) = \left(1/2 + \sum_{k\geq 1} \rho_{k,n} \cos kx\right) \geq 0,$$

P.P. Korovkin [1960] obtained $\rho_{1,n} \to 1$ as the condition for the convergence $L_n f \to f \in C_{2\pi}$ and the following quantitative error estimate:

$$L_n(f;x) - f(x) \leq \left(1 + m\pi[(1-\rho_{1,n})/2]^{1/2}\right)\omega(1/m),$$

with m any positive number and $\omega(\delta)$ the modulus of continuity of the function $f(x)$.

In our treatment below we restrict to the case of linear positive filters. One of the cases, namely that of De La Vallée-Poussin operators V_n, that we take up for a detailed study actually does much more. I.J. Schoenberg [1959] has shown that these operators are variation diminishing, i.e., the number of zero crossings of $V_n(f)$ do not exceed the number of zero crossings of the function f for any $f \in C_{2\pi}$. The convolutions with the Gaussian function $G(X,\sigma)$ are also known to possess this propertry. It is therefore no wonder that these convolutions seem to provide a satisfactory usage in image segmentation.

Let $f \in C_{2\pi}$, the space of 2π-periodic continuous functions and the ordinary modulus of continuity of f be defined by:

$$\omega(f,\delta) = \max_{|x-y|\leq\delta} |f(x) - f(y)|, (\delta > 0).$$

R. Bojanic and O. Shisha [1974] considered the discrete convolutions

$$K_{n,m(n)+2}(f;x) = [2/(m(n)+2)]\sum_{1\leq k\leq m(n)+2} f(t_{k,m(n)+2})\phi(x - t_{k,m(n)+2}),$$

associated with the continuous convolutions: $K_n(f;x) = (1/\pi)\dot{\int}_{[-\pi,\pi]} f(t)\phi_n(x-t)dt,$

with non-negative trigonometric polynomials

$$\phi_n(x) = 1/2 + \sum_{1\leq k\leq m(n)} \rho_{k,n} \cos kx,$$

where the nodes of the discrete convolution, in general, are given by

$$t_{k,l(n)} = x_0 + 2\pi k/l(n), 1 \le k \le l(n).$$

R. Bojanic and O. Shisha [1974] considered the case $x_0 = 0$, and obtained the quantitative error estimate

$$K_{n,m(n)+2}(f;x) - f(x) = (1+\pi)\omega\left(f;[(1-\rho_{1,n})/2]^{1/2}\right).$$

In the following we shall study the more general discretizations

$$K_{n,l(n)}(f;x) = [2/l(n)]\sum_{1\le k\le l(n)} f(t_{k,l(n)})\phi_n(t_{k,l(n)} - x),$$

with x_0 as any fixed point, and $l(n) \ge m(n)+2$ a positive integer. The basic idea behind these discretizations is the trigonometric polynomial precision 'n' of a Gaussian quadrature based on 'n+1' equidistant nodes:

$$\int_{[-\pi,\pi]} T_n(x)dx = 2\pi/(n+1)\sum_{0\le k\le n} T_n(x_k), (x_k = x_0 + 2\pi k/(n+1), 0 \le k \le n),$$

where x_0 is an arbitrary point and $T_n(x)$ is any trigonometric polynomial of order n. Accordingly, if $f(x)$ is a trigonometric polynomial of order p,

$$K_{n,l(n)}(f;x) = K_n(f;x),$$

provided $l(n) \ge m(n) + p$. Using it, a generalization of the R. Bojanic and O. Shisha [1974] $\omega(f;\delta)$ error estimate, and also an $\omega(f';\delta)$ error estimate is:

Lemma 3.2.1.

Let $l(n) \ge m(n)+2$. Then:

(i) $K_{n,l(n)}(f;x) - f(x) \le (1+\pi)\omega(f;[(1-\rho_{1,n})/2]^{1/2}), f \in C_{2\pi}$; and

(ii) $K_{n,l(n)}(f;x) - f(x) \le \pi^2 f'(x)(1-\rho_{1,n})/2 + 2\pi^2[(1-\rho_{1,n})/2]^{1/2}$

$$\times\omega(f';[(1-\rho_{1,n})/2]^{1/2}), f' \in C_{2\pi}.$$

Proof: As the operators $K_{n,l(n)}$ are linear positive (R.A. DeVore [1972 Th. 2.4]), for $f \in C_{2\pi}$,

$$K_{n,l(n)}(f;x) - f(x) \le f(x)\,|1 - K_{n,l(n)}(1;x)| + (K_{n,l(n)}(1;x) + \pi(K_{n,l(n)}(1;x))^{1/2})$$

$$\times\omega(f;K_{n,l(n)}(\sin^2[(t-x)/2];x)),$$

and, for $f \in C_{2\pi}$,

$$K_{n,l(n)}(f;x)-f(x) \le f(x)\,1-K_{n,l(n)}(1;x)+f'(x)\Big(K_{n,l(n)}(\sin(t-x);x)$$
$$+\pi^2 K_{n,l(n)}(\sin^2[(t-x)/2];x)\Big)+\pi^2\Big(1+\big(K_{n,l(n)}(1;x)\big)^{1/2}\Big)$$
$$\times\Big(K_{n,l(n)}(\sin^2[(t-x)/2];x)\Big)^{1/2}\times\omega(f;K_{n,l(n)}(\sin^2[(t-x)/2];x)).$$

Since $\phi_n(t-x)$, $\sin t\,\phi_n(t-x)$, $\cos t\,\phi_n(t-x)$ are trigonometric polynomials of orders $m(n)$, $m(n)+1$, and, $m(n)+1$, respectively, for $l(n)\ge m(n)+2$, we have:

$$K_{n,l(n)}(1;x)=1,\ K_{n,l(n)}(\sin t;x)=\rho_{1,n}\sin x,\ K_{n,l(n)}(\cos t;x)=\rho_{1,n}\cos x.$$

Hence,

$$K_{n,l(n)}(\sin(t-x);x)=K_{n,l(n)}(\sin t;x)\cos x-K_{n,l(n)}(\cos t;x)\sin x=0,\ \text{and}$$

$$K_{n,l(n)}(\sin^2[(t-x)/2];x)=(1/2)[1-K_{n,l(n)}(\cos t;x)\cos x-K_{n,l(n)}(\sin t;x)\sin x]$$
$$=(1-\rho_{1,n})/2,$$

from which the Lemma follows.

Also, one has the following asymptotic error formula in the approximation of $f(x)$ by the discretizations $K_{n,l(n)}(f;x)$:

Theorem 3.2.2.

If $l(n)>m(n)+2$, f is a bounded 2π-periodic function, and f'' exists at a point x,

$$K_{n,l(n)}(f;x)-f(x)=(1-\rho_{1,n})f''(x)+o(1-\rho_{1,n}),$$

iff for some $k\ge 2$, $\lim_{n\to\infty}(1-\rho_{k,n})/(1-\rho_{1,n})=k^2$.

Proof: Both $\sin 2t\,\phi_n(t-x)$ and $\cos 2t\,\phi_n(t-x)$ are trigonometric polynomials of order $m(n)+2$. Hence, if $l(n)>m(n)+2$, also there hold:

$$K_{n,l(n)}(\sin 2t;x)=\rho_{2,n}\sin 2x,\ K_{n,l(n)}(\cos 2t;x)=\rho_{2,n}\cos 2x.$$

From this the result follows since for a sequence L_n of linear positive operators there holds (R.K.S. Rathore [1974a, Th. 1.2.1.5]) the asymptotic relation

$$L_n(f;x)-f(x)=f(x)\{L_n(1;x)-1\}+f'(x)L_n(\sin(t-x);x)$$
$$+[2f''(x)+o(1)]L_n(\sin^2(t-x)/2;x),$$

as $n\to\infty$, iff for some $k=2, 3, \ldots,$

$$\lim_{n\to\infty}(L_n(\sin^2 k(t-x)/2;x))/(L_n(\sin^2(t-x)/2;x))=k^2.$$

It may also be noted that the asymptotic formula holds uniformly in x on each closed sub-interval of an open interval on which f'' exists and is continuous. In particular, for twice continuously differentiable functions in $C_{2\pi}$ the asymptotic formula holds uniformly in x on the real line (R.K.S. Rathore [1974a, Th. 1.2.1.5]).

3.3. DEGREE OF APPROXIMATION BY DISCRETE CONVOLUTIONS

In particular, for the De La Vallée-Poussin, Korovkin, Jackson, and the generalized Jackson continuous convolutions given above, their discretizations are given as follows:

$$V_{n,l(n)}(f;x) = [2/l(n)][(2n)!!/(2(2n-1)!!)]\sum_{1\le k\le l(n)} f(t_{k,l(n)})\cos^{2n}(t_{k,l(n)}-x)/2,,$$

$$A_{n,l(n)}(f;x) = \frac{2}{l(n)}\sum_{1\le k\le l(n)} f(t_{k,l(n)})\left(\frac{1}{n}\right)\left[\frac{\sin(\pi/n)\cos n[(t_{k,l(n)}-x)/2]}{\cos(t_{k,l(n)}-x)-\cos(\pi/n)}\right]^2,$$

$$L_{2n-2,l(n)}(f;x) = [2/l(n)]\sum_{1\le k\le l(n)} f(t_{k,l(n)})[3/(2n(2n^2+1))]\left[\frac{\sin[n(t_{k,l(n)}-x)/2]}{\sin[(t_{k,l(n)}-x)/2]}\right]^4,$$

$$L_{np-p,l(n)}(f;x) = [2/l(n)]\sum_{1\le k\le l(n)} f(t_{k,l(n)})[\pi/A_{np-p}]\left[\frac{\sin[n(t_{k,l(n)}-x)/2]}{\sin[(t_{k,l(n)}-x)/2]}\right]^{2p},(p\ge 3).$$

For these discrete filters, there hold the asymptotic formulae:

$$V_{n,l(n)}(f;x) - f(x) = f''(x)/n + o(1/n), n\to\infty, (l(n) > n+1),$$

$$A_{n-2,l(n)}(f;x) - f(x) = \pi^2 f''(x)/(2n^2) + o(1/n^2), n\to\infty, (l(n) > n),$$

$$L_{2n-2,l(n)}(f;x) - f(x) = 3f''(x)/(2n^2) + o(1/n^2), n\to\infty, (l(n) > 2n),$$

$$L_{np-p,l(n}(f;x) - f(x) = \left(1-\frac{\rho_1^{(np-p)}}{\rho_0^{(np-p)}}\right) f''(x) + o(1/n^2), n\to\infty, (l(n) > np-p+2).$$

These asymptotic formulae are the same as those for the original operators (see e.g., P.P. Korovkin [1960], I.P. Natanson [1964], R.K.S. Rathore [1974a]). To

study the direct and inverse results on their degree of approximation, some definitions needed are as follows: Let $C_{2\pi}$ as usual denote the space of 2π-periodic continuous functions on R, and $C_{2\pi}{}^2$ the sub-space of functions having continuous second order derivative. Let $\|f\| = \max |f(x)|$, $f \in C_{2\pi}$, and let $\|g\|_2 = \|g\| + \|g''\|$, $g \in C_{2\pi}{}^2$. Let the Peetre's K-functional for $f \in C_{2\pi}$ be defined as:

$$K(f;t) = \inf\{\|f - g\| + t^2 \|g''\| : g \in C_{2\pi}{}^2\}, t > 0.$$

Lemma 3.3.1.

If $l(n) \geq m(n) + 2$, for some constant M,

$$\|g - K_{n,l(n)} g\| \leq M(1 - \rho_{1,n}) \|g''\|_2, g \in C_{2\pi}{}^2.$$

Proof: By $\omega(f';\delta)$ error estimate,

$$K_{n,l(n)}(g;x) - g(x) \leq \pi^2 |g'(x)|(1 - \rho_{1,n})/2 + 2\pi^2[(1 - \rho_{1,n})/2]^{1/2} \times \omega(g';[(1 - \rho_{1,n})/2]^{1/2}) \leq ((1 - \rho_{1,n})/2)(\pi^3 + 2\pi^2)\|g''\|,$$

from which the result follows.

Lemma 3.3.2.

If $l(n) \geq m(n) + 2$,

$$\|K_{n,l(n)}''(g;x)\| \leq (\pi m(n)/2)^2 (1 - \rho_{1,n}) \|g''\|_2, g \in C_{2\pi}{}^2.$$

Proof: Choose α such that $g'(\alpha) = 0$. Then,

$$K_{n,l(n)}(g;x) = g(\alpha) + K_{n,l(n)}((t - x)^2 g''(\xi)/2;x),$$

so that by the Bernstein's inequality,

$$\begin{aligned} \|K_{n,l(n)}''(g;x)\| &= \|K_{n,l(n)}''((t - x)^2 g''(\xi)/2;x)\| \\ &\leq [m(n)]^2 \|K_{n,l(n)}((t - x)^2 g''(\xi)/2;x)\| \\ &\leq [\pi m(n)/2]^2 \|K_{n,l(n)}(2\sin^2(t - x)/2;x)\| \|g''\|_2 \\ &\leq [\pi m(n)/2]^2 (1 - \rho_{1,n}) \|g''\|_2. \end{aligned}$$

Theorem 3.3.3.

Let $f(x) \in C_{2\pi}$, and $\psi \in S$. If $1-\rho_{1,n} = O([m(n)]^{-2}) \sim 1/n^2$ and $l(n) \geq m(n)+2$, then:

(i) $K_{n,l(n)}f - f \leq AK(f;(1-\rho_{1,n})^{1/2}) \to 0, n \to \infty$, where A is a constant;

(ii) $K_{n,l(n)}f - f = O(\psi(n^{-2})), n \to \infty$, iff $K(f;t) = O(\psi(t^2)), t \to 0$;

(iii) $K_{n,l(n)}f - f = o(\psi(n^{-2})), n \to \infty$, iff $K(f;t) = o(\psi(t^2)), t \to 0$;

(iv) $K_{n,l(n)}f - f \sim K(f;n^{-1})$, if $K(f;t^{1/2}) \in S$.

Proof: For (i) we have

$$\begin{aligned} K_{n,l(n)}f - f &\leq g - f + K_{n,l(n)}g - g + K_{n,l(n)}(f-g) \\ &\leq 2\,\|f-g\| + M(1-\rho_{1,n})\,\|g''\|. \end{aligned}$$

Taking infimum over $g \in C_{2\pi}^2$,

$$K_{n,l(n)}f - f \leq AK(f;(1-\rho_{1,n})^{1/2}),$$

where $A = \max\{2, M\}$.

To prove (ii)-(iv), using (i), if $K(f;t) = O(\psi(t^2)), t \to 0$,

$$K_{n,l(n)}f - f = O(\psi(1-\rho_{1,n})), n \to \infty,$$

and, if, $K(f;t) = o(\psi(t^2)), t \to 0$,

$$K_{n,l(n)}f - f = o(\psi(1-\rho_{1,n})), n \to \infty.$$

For the inverse parts of (ii)- (iv), using the previous two Lemmas,

$$\begin{aligned} K(f;t) &\leq K(f - K_{n,l(n)}(f;t)) + K(K_{n,l(n)}(f-g;t)) + K(K_{n,l(n)}(g;t)) \\ &\leq \|f - K_{n,l(n)}f\| + t^2\left\{ \|K_{n,l(n)}''(f-g)\| + \|K_{n,l(n)}''g\| \right\} \\ &\leq \|f - K_{n,l(n)}f\| + t^2(\pi m(n)/2)^2\left\{ \|f-g\| + [(1-\rho_{1,n})]^2\,\|g''\| \right\}. \end{aligned}$$

Taking infimum over $g \in C_{2\pi}^2$,

$$K(f;t) \leq M\left\{ \|f - K_{n,l(n)}f\| + (tn)^2[K(f;n^{-1})] \right\},$$

and the results follow from Lemma 1.2.2.

For the familiar Bernsteinian orders $n^{-\alpha}$, $0<\alpha<2$, we have the immediate:

Corollary 3.3.4.

Let $f(x)\in C_{2\pi}$, $0<\alpha<2$, $1-\rho_{1,n}=O([m(n)]^{-2})\sim 1/n^2$ and $l(n)\geq m(n)+2$. Then:

(i) $K_{n,l(n)}f-f\ =O(n^{-\alpha}), n\to\infty$, iff $K(f;t)=O(t^{\alpha}), t\to 0$;

(ii) $K_{n,l(n)}f-f\ =o(n^{-\alpha}), n\to\infty$, iff $K(f;t)=o(t^{\alpha}), t\to 0$;

(iii) $K_{n,l(n)}f-f\ \sim n^{-\alpha}$, if $K(f;t^{1/2})\sim n^{-\alpha}$.

For the De La Vallée-Poussin kernel there holds: $1-\rho_{1,n}=O(1/n)$. Hence the above results are not applicable to the discrete filters $V_{n,l(n)}(f;x)$, which we shall take up separately in the next section for a much more detailed study.

However, since the relation $1-\rho_{1,n}=O([m(n)]^{-2})\sim 1/n^2$ is valid for the earlier mentioned Korovkin, the generalized Jackson, and Jackson kernels (the case p = 2, in the following), for the discrete convolutions $A_{n-2,l(n)}(f;x)$, $L_{np-p,l(n)}(f;x)$ and $L_{2n-2,l(n)}(f;x)$, we obtain the following:

Corollary 3.3.5.

Let $f(x)\in C_{2\pi}$, and $\psi\in S$. Then, if $l(n)\geq n$:

(i) $A_{n-2,l(n)}f-f\ \leq A\omega_2(f;1/n)\to 0, n\to\infty$, for a constant A;

(ii) $A_{n-2,l(n)}f-f\ =O(\psi(n^{-2})), n\to\infty$, iff $\omega_2(f;t)=O(\psi(t^2)), t\to 0$;

(iii) $A_{n-2,l(n)}f-f\ =o(\psi(n^{-2})), n\to\infty$, iff $\omega_2(f;t)=o(\psi(t^2)), t\to 0$;

(iv) $A_{n-2,l(n)}f-f\ \sim\omega_2(f;n^{-1})$, if $\omega_2(f;t^{1/2})\in S$.

Corollary 3.3.6.

Let $f(x)\in C_{2\pi}$, $p\geq 2$, and $\psi\in S$. Then, if $l(n)\geq np-p+2$:

(i) $L_{np-p,l(n)}f-f\ \leq A\omega_2(f;1/n)\to 0, n\to\infty$, for a constant A;

(ii) $L_{np-p,l(n)}f-f\ =O(\psi(n^{-2})), n\to\infty$, iff $\omega_2(f;t)=O(\psi(t^2)), t\to 0$;

(iii) $L_{np-p,l(n)}f - f = o(\psi(n^{-2})), n \to \infty$, iff $\omega_2(f;t) = o(\psi(t^2)), t \to 0$;

(iv) $L_{np-p,l(n)}f - f \sim \omega_2(f;n^{-1})$, if $\omega_2(f;t^{1/2}) \in S$.

In particular for the Bernsteinian orders $n^{-\alpha}$, $0 < \alpha < 2$, there hold:

Corollary 3.3.7.

Let $f(x) \in C_{2\pi}$, $0 < \alpha < 2$, and $l(n) \geq n$. Then:

(i) $A_{n-2,l(n)}f - f = O(n^{-\alpha}), n \to \infty$, iff $\omega_2(f;t) = O(t^{\alpha}), t \to 0$;

(ii) $A_{n-2,l(n)}f - f = o(n^{-\alpha}), n \to \infty$, iff $\omega_2(f;t) = o(t^{\alpha}), t \to 0$;

(iii) $A_{n-2,l(n)}f - f \sim n^{-\alpha}$, if $\omega_2(f;t) \sim t^{\alpha}$.

Corollary 3.3.8.

Let $f(x) \in C_{2\pi}$, $0 < \alpha < 2$, $p \geq 2$, and $l(n) \geq np - p + 2$. Then:

(i) $L_{np-p,l(n)}f - f = O(n^{-\alpha}), n \to \infty$, iff $\omega_2(f;t) = O(t^{\alpha}), t \to 0$;

(ii) $L_{np-p,l(n)}f - f = o(n^{-\alpha}), n \to \infty$, iff $\omega_2(f;t) = o(t^{\alpha}), t \to 0$;

(iii) $L_{np-p,l(n)}f - f \sim n^{-\alpha}$, if $\omega_2(f;t) \sim t^{\alpha}$.

In these corollaries, we have used the well-known equivalence

$$AK(f;t) \leq \omega_2(f;t) \leq BK(f;t), (t > 0)$$

of $K(f;t)$ and $\omega_2(f;t)$, where A and B are positive constants, and indeed the second order modulus of smoothness is defined as

$$\omega_2(f;t) = \sup_{|h| \leq t} \left(\sup_{x \in [-\pi,\pi]} | f(x+h) - 2f(x) + f(x-h) | \right), f \in C_{2\pi}.$$

3.4. DISCRETE CONVOLUTIONS WITH DE LA VALLÉE-POUSSIN KERNELS

In this section we study discrete convolutions with the De La Vallée-Poussin kernels, whose continuous version is given by:

$$V_n(x) = V_n(f;x) = \frac{(2n)!!}{2\pi(2n-1)!!}\int_{-\pi}^{\pi} f(t)\cos^{2n}\left(\frac{t-x}{2}\right)dt..$$

Some well known results (I.P. Natanson [1964]) about the continuous convolutions $V_n(f;x)$ are:

(i) If a function $f(x)\in C_{2\pi}$ has modulus of continuity $\omega(\delta)$, then the inequality

$$V_n(x) - f(x) \le 3\omega\left(\frac{1}{\sqrt{n}}\right)$$

holds for all values of x.

(ii) If $U_n(\alpha) = \sup\{\max|V_n(x) - f(x)|\}$, wherein $f(x)$ runs through all functions with period 2π of class $\mathrm{Lip}_1\alpha$, then

$$U_n(\alpha) = \frac{2^\alpha}{\sqrt{\pi}\,n^\alpha}\,\Gamma\left(\frac{1+\alpha}{2}\right) + \frac{\rho_n}{\sqrt{n^\alpha}},$$

where $\lim_{n\to\infty}\rho_n = 0$.

(iii) If the function $f(x)\in C_{2\pi}$ possesses for a value of x a finite second derivative $f''(x)$, then the formula

$$V_n(x) = f(x) + \frac{f''(x)}{n} + \frac{\rho_n}{n}$$

with $\lim_{n\to\infty}\rho_n = 0$ holds true for this x.

(iv) If at a point x, $f(x)\in C_{2\pi}$ has a finite derivative $f'(x)$, then the relation

$$\lim_{n\to\infty} V_n'(x) = f'(x)$$

holds true for this value of x.

Generalizations of some of the above results by P.C. Sikkema and R.K.S. Rathore [1976] are:

(v) Let $0<\alpha\le 1$, and p be a positive integer. If

$$E(\alpha, p; V_n, x) = \sup_f\left\{\left|V_m\left(f(t) - \sum_{\nu=0}^{p}\frac{f^{(\nu)}(x)}{\nu!}(t-x)^\nu;x\right)\right|\right\},\quad x\in\mathrm{R},$$

where f runs through all functions with $f^{(p)} \in \mathrm{Lip}_1\alpha$, then with $(a)_p = \prod_{i=1}^{p}(a+i-1)$,

$$\frac{2^{p+2\alpha-1}\Gamma(\frac{p+\alpha+1}{2})}{(1+\alpha)_p\sqrt{\pi}} \le \lim_{n\to\infty} n^{\frac{p+\alpha}{2}} E(\alpha, p; V_n, x) \le \frac{2^{p+\alpha}\Gamma(\frac{p+\alpha+1}{2})}{(1+\alpha)_p\sqrt{\pi}}.$$

(vi) If p is a positive integer, then

$$\lim_{n\to\infty} n^{\frac{p+1}{2}} E(1, p; V_n, x) = \frac{2^{p+1}\Gamma(\frac{p}{2}+1)}{(p+1)!\sqrt{\pi}}.$$

(vii) Let m be a positive integer. If $f \in C_{2\pi}$ and at a point $x \in \mathrm{R}$, $f^{(m)}(x)$ exists, then

$$\lim_{n\to\infty} \frac{d^m}{dx^m} V_n(f; x) = f^{(m)}(x).$$

(viii) $V_n(\sin^{2k} t/2; 0) \cong \pi^{-1/2}\Gamma\left(\frac{2k+1}{2}\right) n^{-k}$, $k \ge 1$. (V.S.N. Kaliprasad [2000 Lemma 3.2.1.])

Lemma 3.4.1.

Let $m \ge 0$, $n \ge 1$ be positive integers, and $D \equiv d/dx$. Then

$$D^m\left(\cos^{2n}(x/2)\right) = \left(\cos^{2n}(x/2)\right)\left(\sum_{0\le 2j\le[m/2]} P_{m,m-j}(n)\tan^{m-2j}(x/2)\right),$$

where $[m/2]$ is the integral part of $m/2$, and $P_{m,j}(n)$, $(0 \le j \le [m/2])$, is an algebraic polynomial in n of degree $\le j$.

Proof: For $m = 0$, the result is trivial, with $P_{m,0}(n) = 1$. Since,

$$D\left(\cos^{2n}(x/2)\right) = \left(\cos^{2n}(x/2)\right)(-n)\tan(x/2),$$

and

$$D^2[\cos^{2n}(x/2)] = [\cos^{2n}(x/2)][n(n-1/2)\tan^2(x/2) - 2n],$$

the result is true for $m = 1, 2$. Hence assuming it for m,

$$\begin{aligned} D^{m+1}\left(\cos^{2n}(x/2)\right) &= \left(\sum_{0\le j\le[m/2]} P_{m,m-j}(n) D\left\{\cos^{2n}(x/2)\tan^{m-2j}(x/2)\right\}\right) \\ &= \cos^{2n}(x/2)\sum_{0\le j\le[m/2]} P_{m,m-j}(n)\left\{(m/2-j)-n\right)\tan^{m+1-2j}(x/2) \\ &\qquad + (m/2-j)\tan^{m+1-2(j+1)}(x/2) \end{aligned}$$

$$= (\cos^{2n}(x/2))\{\sum_{0\le j\le[(m+1)/2]}[P_{m,m-j}(n)((m/2-j)-n) + (m/2-j+1)P_{m,m-j+1}(n)]\tan^{m+1-2j}(x/2)\},$$

and $P_{m+1,m+1-j}(n) = \left(P_{m,m-j}(n)((m/2-j)-n) + (m/2-j+1)P_{m,m-j+1}\right)$ is a polynomial of degree $\le m+1-j$. Hence the result holds for $m+1$, completing the proof.

Lemma 3.4.2.

There exist trigonometric polynomials $T_k(x)$, $0 \le k \le 2p$, of order p such that

$$T_k(x) = x^k + O(x^{2p+1}),\ 0 \le k \le 2p,\ p \ge 1.$$

Proof: Let $\Omega^{(k)} \equiv \Omega_{2p}{}^{(k)}$ be the functions defined by (R.K.S. Rathore [1974b])

$$\Omega^{(k)}(t,x) = -k!\Delta^{(k)}(t,x)/\Delta^{(0)}(x,x),$$

where $\Delta^{(k)}(t,x)$, $k = 0, 1, \ldots, 2p$, are the determinants of order $(2p+1)$ given by $\Delta^{(k)}(t,x) = \left(a_{ij}^{k}\right)_{i,j=1}^{m+1}$, where with $[j/2]$ denoting the largest integer not greater than $j/2$,

$$a_{1j}^{(k)} = \begin{cases} 1, & j=1 \\ \sin[j/2]t, & j \text{ is even}, \\ \cos[j/2]t, & j \text{ is odd} \end{cases}$$

$$a_{k+1\,j}^{(k)} = \begin{cases} 1, & j=1 \\ \sin[j/2]x, & j \text{ is even}, \\ \cos[j/2]x, & j \text{ is odd} \end{cases} \qquad k \ne 0,$$

and

$$a_{ij}^{(k)} = \frac{d^{i-1}}{dx^{i-1}} a_{k+1\,j}^{(k)},\ 1 \ne i \ne k+1.$$

Since the functions $\Omega^{(k)}(t,x)$ satisfy

$$\Omega^{(k)}(t,x) = (t-x)^k + O(|t-x|^{2p+1}),$$

taking $T_k(x) = T_{k,2p}(x) = \Omega^{(k)}(x,0)$, the result follows.

In conjunction with Taylor's expansion, as an immediate consequence of the Lemma, we have the following:

Corollary 3.4.3.

If at a point 'x', $f^{(m)}(x)$ exists where $f(t)$ is a bounded 2π-periodic function, then:

$$f(t) = f(x) + \sum_{1\le k\le m}\left(f^{(k)}(x)/k!\right)T_k(t-x) + h_x(t)(t-x)^m, \qquad 1\le m\le 2p,$$

where $h_x(t)$, with $h_x(x)=0$, is bounded and is continuous at $t = x$. Moreover, if $f^{(m)}(x)$ exists at all 'x' and belongs to $C_{2\pi}$, for an any $\varepsilon > 0$, there exists a $\delta > 0$, independent of x, such that $h_x(t) < \varepsilon$, $0 < |t\text{-}x| < \delta$.

The next result establishes the basic approximation of pointwise derivatives by the derivatives of the discrete De La Vallée-Poussin operators $V_n(f;x)$:

Theorem 3.4.5. (Approximation of derivatives).

If f is a bounded 2π-periodic function such that for some natural number m, and at some point x, $f^{(m)}(x)$ exists, then with $l(n) > n+[m/2]+1$:

$$\lim_{n\to\infty} V_{n,l(n)}{}^{(m)}(f;x) = f^{(m)}(x).$$

Moreover, if $f^{(m)}(x)$ exists for all x and belongs to $C_{2\pi}$, the convergence is uniform in x.

Proof: Using the Corollary above, where we choose $2p = m$, if m is even, and $2p = m+1$, if m is odd, we have:

$$V_{n,l(n)}{}^{(m)}(f;x) = V_{n,l(n)}{}^{(m)}\left(f(x) + \sum_{1\le j\le m}(j!)^{-1}f^{(j)}(x)T_j(t-x) + o((x-t)^m;x)\right)$$

$$= \left[(2n)!!/(l(n)(2n-1)!!)\right]\sum_{1\le k\le l(n)}\left(\sum_{0\le j\le m}f^{(j)}(x)T_j(t_{k,l(n)}-x) + o(x-t_{k,l(n)})^m\right)$$

$$\times \cos^{2n}(x-t_{k,l(n)})/2\sum_{0\le 2j\le[m/2]}P_{m,m-j}(n)\tan^{m-2j}((x-t_{k,l(n)})/2)$$

$$= \sum_{1\le j\le m}f^{(j)}(x)V_{n,l(n)}{}^{(m)}(T_j(t-x);x) + \left[(2n)!!/(l(n)(2n-1)!!)\right]\sum_{1\le k\le l(n)}o(x-t_{k,l(n)})^m$$

$$\times \cos^{2n}(x-t_{k,l(n)})/2\sum_{0\le 2j\le[m/2]}P_{m,m-j}(n)\tan^{m-2j}((x-t_{k,l(n)})/2).$$

Since T_j's are fixed trigonometric polynomials satisfying $T_j{}^m(0) = m!\delta_{jm}$, $0\le j\le m\le 2p$, where δ_{jm} is a Kronecker delta,

$$\sum_{1\le j\le m}(j!)^{-1}f^{(j)}(x)V_{n,l(n)}{}^{(m)}(T_j(t-x);x) \to f^{(m)}(x), n\to\infty, (l(n) > n+p).$$

Hence, it remains to show that the contribution of the o-term tends to zero. For this, in view of the Corollary above, given an arbitrary $\varepsilon > 0$, we can find a $\delta > 0$, such that,

$$o(x-t_{k,l(n)})^m \le \varepsilon\, x-t_{k,l(n)}{}^m + (A/\delta^2)\, x-t_{k,l(n)}{}^{m+2},$$

for some constant A, and for all $t_{k,l(n)}$. Hence, with $\mu = m$, if m is even, and $\mu = m+1$, if m is odd, in view of the 2π-periodicity of $f(t)$ using which we could assume that $(x-t_{k,l(n)}) \le \pi$ we have, for $l(n) > n\text{-}\mu/2+m+1 = n+[m/2]+1$,

$$\begin{aligned}
&[(2n)!!/(l(n)(2n-1)!!)]\textstyle\sum_{1\le k\le l(n)} o(x-t_{k,l(n)})^m \cos^{2n}(x-t_{k,l(n)})/2 \\
&\qquad \times \textstyle\sum_{0\le 2j\le\lfloor m/2\rfloor} P_{m,m-j}(n)\tan^{m-2j}((x-t_{k,l(n)})/2) \\
&= O\Big(\textstyle\sum_{0\le 2j\le[m/2]} P_{m,m-j}(n)\, V_{n-\mu/2+j,l(n)}\big(\varepsilon\sin^{2m-2j}((x-t_{k-l(n)})/2) \\
&\qquad\qquad + (A/\delta^2)\sin^{2m+2-2j}((x-t_{k-l(n)})/2);x\big)\Big) \\
&= O\Big(\textstyle\sum_{0\le 2j\le[m/2]} n^{m-j} V_{n-\mu/2+j}\big(\varepsilon\sin^{2m-2j}(t/2) + (A/\delta^2)\sin^{2m+2-2j}(t/2);0\big)\Big) \\
&= O\Big(\textstyle\sum_{0\le 2j\le[m/2]} n^{m-j}\big(\varepsilon n^{-m+j} (A/\delta^2) n^{-m-1+2j}\big)\Big),
\end{aligned}$$

by (viii) above, from which, due to the arbitrariness of $\varepsilon > 0$, it follows that the contribution of the o-term tends to zero as $n \to \infty$.

The uniformity of the convergence:

$$\lim_{n\to\infty} V_{n,l(n)}{}^{(m)}(f;x) = f^{(m)}(x),$$

in the case when $f^{(m)}(x)$ exists for all x and belongs to $C_{2\pi}$, is clear since then A and the δ in the above could be chosen to be independent of x.

The next result provides an asymptotic formula of Voronovskaya type in the approximation of a derivative at a point:

Theorem 3.4.6. (Voronovskaya Type Asymptotic Formula).

If f is a bounded 2π-periodic function such that for a natural number m, at some point x, $f^{(m+2)}(x)$ exists, then with $l(n) > n+[m/2]+2$:

$$\lim_{n\to\infty} n\big(V_{n,l(n)}{}^{(m)}(f;x) - f^{(m)}(x)\big) = f^{(m+2)}(x).$$

Moreover, if $f^{(m+2)}(x)$ exists for all x and belongs to $C_{2\pi}$, the convergence is uniform in x.

Proof: Using the Corollary above, where we choose $p = [(m+3)/2]$, i.e., $2p = m+2$, if m is even, and $2p = m+3$, if m is odd, we have:

$$V_{n,l(n)}^{(m)}(f;x) = V_{n,l(n)}^{(m)}\Big(f(x) + \sum_{1\le j\le m+2}(j!)^{-1} f^{(j)}(x)T_j(t-x) + o((x-t)^{m+2};x)\Big)$$

$$= [(2n)!!/(l(n)(2n-1)!!)]\sum_{1\le k\le l(n)}\Big(\sum_{0\le j\le m+2}(j!)^{-1} f^{(j)}(x)T_j(t_{k,l(n)}-x)$$

$$+ o(x-t_{k,l(n)})^{m+2}\Big)\times \cos^{2n}(x-t_{k,l(n)})/2\Big)$$

$$\times \sum_{0\le 2j\le [m/2]} P_{m,m-j}(n)\tan^{m-2j}((x-t_{k,l(n)})/2)\Big)$$

$$= \sum_{1\le j\le m+2}(j!)^{-1} f^{(j)}(x)V_{n,l(n)}^{(m)}(T_j(t-x);x) + [(2n)!!/(l(n)(2n-1)!!)]$$

$$\times \sum_{1\le k\le l(n)} o(x-t_{k,l(n)})^{m+2}\Big)\cos^{2n}(x-t_{k,l(n)})/2\Big)$$

$$\times \sum_{0\le 2j\le [m/2]} P_{m,m-j}(n)\tan^{m-2j}((x-t_{k,l(n)})/2)\Big).$$

First of all we consider the contribution of the o-term, for which given an arbitrary $\varepsilon > 0$, we can find a $\delta > 0$, such that,

$$o(x-t_{k,l(n)})^{m+2} \le \varepsilon\, |x-t_{k,l(n)}|^{m+2} + (B/\delta^2)\,|x-t_{k,l(n)}|^{m+4},$$

for some constant B, and for all $t_{k,l(n)}$. Hence, with $\mu = 2[(m+1)/2]$, i.e., $\mu = m$, if m is even, and $\mu = m+1$, if m is odd , in view of the 2π-periodicity of $f(t)$ we could assume that $(x-t_{k,l(n)}) \le \pi$, so that $|x-t_{k,l(n)}| \le \pi \sin((x-t_{k,l(n)})/2)$. Hence, for $l(n) > n-\mu/2+m+2 = n+[m/2]+2$:

$$[(2n)!!/(l(n)(2n-1)!!)]\sum_{1\le k\le l(n)} o(x-t_{k,l(n)})^{m+2}\Big)\cos^{2n}(x-t_{k,l(n)})/2\Big)$$

$$\times \sum_{0\le 2j\le [m/2]} P_{m,m-j}(n)\tan^{m-2j}((x-t_{k,l(n)})/2)$$

$$= O\Big(\sum_{0\le 2j\le [m/2]} P_{m,m-j}(n)$$

$$\times V_{n-\mu/2+j,l(n)}\Big(\varepsilon \sin^{2m+2-2j}((x-t_{k-l(n)})/2) + (B/\delta^2)\sin^{2m+4-2j}((x-t_{k-l(n)})/2);x\Big)\Big)$$

$$= O\Big(\sum_{0\le 2j\le [m/2]} n^{m-j} V_{n-\mu/2+j}\Big(\varepsilon \sin^{2m+2-2j}(t/2) + (B/\delta^2)\sin^{2m+4-2j}(t/2);0\Big)\Big)$$

$$= O\Big(\sum_{0\le 2j\le [m/2]} n^{m-j}\Big(\varepsilon n^{-m-1+j} + (B/\delta^2)n^{-m-2+2j}\Big)\Big),$$

by (viii) above, from which, due to the arbitrariness of $\varepsilon > 0$, it follows that the contribution of the o-term is $o(n^{-1})$, as $n \to \infty$.

Next, using [.] for the integral part, since T_j's are fixed polynomials of order $p = [(m+3)/2]$ and $p+n = [(m+3)/2]+n \le n+[m/2]+2 < l(n)$,

$$V_{n,l(n)}(T_j(t-x);x) = V_n(T_j(t-x);x) = V_n(T_j;0).$$

Let $T_j(x) = a_0/2 + \sum_{1\le k\le p}\{a_k \cos kx + b_k \sin kx\}$. In view of the Voronovskaya asymptotic formula for the operators V_n,

$$V_n(T_j;x) = a_0/2 + \sum_{1\le k\le p} \rho_{k,n}[a_k \cos kx + b_k \sin kx]$$
$$= a_0/2 + \sum_{1\le k\le p}(1-k^2/n + o(1/n))[a_k \cos kx + b_k \sin kx],$$

so that

$$V_{n,l(n)}{}^{(m)}(T_j;x) = V_n{}^{(m)}(T_j(t-x);x)$$
$$= \left\{a_0/2 + \sum_{1\le k\le p}(1-k^2/n + o(1/n))[a_k \cos kz + b_k \sin kz]\right\}^{(m)}_{z=0}$$
$$= T_j{}^{(m)}(0) - \left\{\sum_{1\le k\le p}(k^2/n)[a_k \cos kz + b_k \sin kz]\right\}^{(m+2)}_{z=0} + o(1/n)$$
$$= T_j{}^{(m)}(0) + (1/n)T_j{}^{(m+2)}(0) + o(1/n).$$

Hence, since $T_j{}^m(0) = m!\delta_{jm}$, $0 \le j \le m \le 2p$, where δ_{jm} is a Kronecker delta, we have

$$\sum_{1\le j\le m+2}(j!)^{-1} f^{(j)}(x) V_{n,l(n)}{}^{(m)}(T_j;x)$$
$$= \sum_{1\le j\le m+2}(j!)^{-1} f^{(j)}(x)\{T_j{}^{(m)}(0) + (1/n)T_j{}^{(m+2)}(0)\} + o(1/n)$$
$$= f^{(m)}(x) + (1/n) f^{(m+2)}(x) + o(1/n),$$

from which the required asymptotic relation

$$V_{n,l(n)}{}^{(m)}(f;x) = f^{(m)}(x) + (1/n) f^{(m+2)}(x) + o(1/n),$$

follows. Noting that if $f^{(m+2)}(x) \in C_{2\pi}$, the δ in the above could be chosen independently of x and, furthermore, that the magnitudes of all the derivatives $f^{(j)}(x)$, $1 \le j \le m+2$, possess a common upper bound, the uniformity part follows, completing the proof of the Theorem.

Let $C_{2\pi}{}^m$, $(m \ge 0)$, denote the space of 2π-periodic functions having continuous m-th order derivative on the real line. Let $\|f\| = \max |f(x)|$, $f \in C_{2\pi}$. Let a Peetre's K-functional, related with the pair $(C_{2\pi}{}^m, C_{2\pi}{}^{m+2})$, for $f \in C_{2\pi}{}^m$, be defined as:

$$K_{m,m+2}(f;t) = \inf\left\{ \|f^{(m)} - g^{(m)}\| + t^2 \|g^{(m+2)}\| : g \in C_{2\pi}{}^{m+2} \right\}, t > 0.$$

In order to handle the approximation of functions as well as their derivatives simultaneously, we would naturally identify $f^{(0)}(x)$ with $f(x)$, and, $C_{2\pi}{}^{0}$ with $C_{2\pi}$.

Lemma 3.4.7.

If $l(n) > n+[m/2]+2$,

$$g^{(m)} - V_{n,l(n)}{}^{(m)} g \le n^{-1} M \; g^{(m+2)} \;, g \in C_{2\pi}{}^{m+2}, \; (m \ge 0),$$

where M is a constant.

Proof: By mean value theorem, for some ξ lying between t and x,

$$g(t) = g(x) + \sum_{1 \le j \le m+1} \left(g^{(j)}(x)/\, j!\right)(t-x)^j + \left(g^{(m+2)}(\xi)/(m+2)!\right)(t-x)^{m+2}, |\, t-x \,| \le \pi.$$

In the rest of the proof, using 2π-periodicity of g and $V_{n,l(n)}g$ we restrict x to belong to $[-\pi, \pi]$, and shall regard the functions $(t-x)^j$ defined by the corresponding value $(t-x)$ for $|\, t-x \,| \le \pi$. Using $p = [(m+3)/2]$, since

$$T_j(x) = x^j + \mathrm{O}(x^{2p+1}), \; 0 \le j \le 2p, \; p \ge 1,$$

we have

$$\begin{aligned}(t-x)^j &= T_j(t-x) + \mathrm{O}((t-x)^{2p+1}) \\ &= T_j(t-x) + \mathrm{O}\left(\sin^{2p}((t-x)/2)\right) \\ &= T_j(t-x) + \mathrm{O}\left(\sin^{m+2}((t-x)/2)\right),\end{aligned}$$

where the O-terms holds uniformly in $|t-x| \le \pi$, so that in view of $\|h'\| \le \pi\|h''\|$, $h \in C_{2\pi}{}^{2}$,

$$g(t) = g(x) + \sum_{1 \le j \le m+1} g^{(j)}(x)/\, j!\, T_j(t-x) + g^{(m+2)}\; \mathrm{O}\left(\sin^{m+2}((t-x)/2)\right), |\, t-x \,| \le \pi.$$

Using the estimates: $V_{n,l(n)}{}^{(m)}(T_j(t-x);x) = j!\delta_{jm} + \mathrm{o}(1/n), \;\; 0 \le j \le m+1,$

in which the $\mathrm{o}(1/n)$ term holds uniformly in x, of the previous Theorem, with $\mu = 2[(m+1)/2]$,

$$\begin{aligned}&V_{n,l(n)}{}^{(m)}(g;x) - g^{(m)}(x) \\ &= V_{n,l(n)}{}^{(m)}\left(g(x) + \sum_{1 \le j \le m+1} (j!)^{-1} g^{(j)}(x) T_j(t-x) + \; g^{(m+2)}\; \mathrm{O}\left(\sin^{m+2}((t-x)/2)\right); x\right) \\ &\qquad - g^{(m)}(x)\end{aligned}$$

$$= g^{(m+2)}\ o(n^{-1}) + V_{n,l(n)}{}^{(m)}\left(O\left(\sin^{m+2}((t-x)/2);x\right)\right)$$

$$= g^{(m+2)}\ \Big(o(n^{-1}) + ((2n)!!/(l(n)(2n-1)!!))\sum_{1\le k\le l(n)} O\left(\sin^{m+2}((x-t_{k,l(n)})/2)\right)$$

$$\times\left(\cos^{2n}(x-t_{k,l(n)})/2\right)\sum_{0\le 2j\le[m/2]} P_{m,m-j}(n)\tan^{m-2j}(x-t_{k,l(n)})/2\Big)$$

$$= g^{(m+2)}\ \left(o(n^{-1}) + O\left(\sum_{0\le 2j\le[m/2]} P_{m,m-j}(n)\ V_{n-\mu/2+j,l(n)}\left(\sin^{2m+2-2j}((x-t_{k,l(n)})/2);x\right)\right)\right)$$

$$= g^{(m+2)}\ \left(o(n^{-1}) + O\left(\sum_{0\le 2j\le[m/2]} n^{m-j}(n^{-m-1+j})\right)\right)$$

$$\le n^{-1}M\ g^{(m+2)},$$

for some constant M, from which the result follows after taking supremum over x.

Lemma 3.4.8.

If $l(n) > n+[q/2]+1$, $V_{n,l(n)}{}^{(q)} g \le K_q\ g^{(q)}$, $g \in C_{2\pi}{}^{q}$, $(q \ge 0)$, K_q being a constant.

Proof: Let $\mu = 2[(q+1)/2]$, and $l(n) > n+[q/2]+1$. For T_j's with $p = \mu/2$, for $q > 0$, using $V_{n,l(n)}{}^{(q)}(1;x) = 1$, and $V_{n,l(n)}{}^{(q)}(T_j(t-x);x) = O(1)$, uniformly in x, we have

$$V_{n,l(n)}{}^{(q)}(g;x)$$

$$= V_{n,l(n)}{}^{(q)}\left(g(x) + \sum_{1\le j\le q-1}(j!)^{-1} g^{(j)}(x)T_j(t-x) + g^{(q)}\ O\left(\sin^{q}((t-x)/2)\right);x\right)$$

$$= g^{(q)}\ O(1) + V_{n,l(n)}{}^{(q)}\left(O\left(\sin^{q}((t-x)/2)\right);x\right)$$

$$= g^{(q)}\ \Big(O(1) + ((2n)!!/(l(n)(2n-1)!!))\sum_{1\le k\le l(n)} O\left(\sin^{q}((x-t_{k,l(n)})/2)\right)$$

$$\times\left(\cos^{2n}(x-t_{k,l(n)})/2\right)\sum_{0\le 2j\le[q/2]} P_{q,q-j}(n)\tan^{q-2j}(x-t_{k,l(n)})/2\Big)$$

$$= g^{(q)}\ \left(O(1) + O\left(\sum_{0\le 2j\le[q/2]} P_{q,q-j}(n)V_{n-\mu/2+j,l(n)}\left(\sin^{2q-2j}((x-t_{k,l(n)})/2)\right);x\right)\right)$$

$$= g^{(q)}\ \left(O(1) + O\left(\sum_{0\le 2j\le[q/2]} n^{q-j}(n^{-q+j})\right)\right)$$

$$\le K_q\ g^{(q)},$$

for some constant K_q. Hence the result, it being trivial for $q = 0$.

Lemma 3.4.9.

If $l(n) > n+[m/2]+2$, $V_{n,l(n)}{}^{(m+2)} h \le nL\ h^{(m)}$, $h \in C_{2\pi}{}^{(m)}$, $m \ge 0$, L being a constant.

Proof: For $l(n) > n+[m/2]+2$, as in the proof of the previous Lemma,

$$
\begin{aligned}
&V_{n,l(n)}{}^{(m+2)}(h;x)\\
&= V_{n,l(n)}{}^{(m+2)}\left(h(x)+\sum_{1\le j\le m-1}(j!)^{-1}h^{(j)}(x)T_j(t-x)+\ h^{(m)}\ \mathrm{O}\left(\sin^m((t-x)/2)\right);x\right)\\
&= h^{(m)}\ \ o(n^{-1})+V_{n,l(n)}{}^{(m+2)}\left(\mathrm{O}\left(\sin^m((t-x)/2)\right),x\right)\\
&= h^{(m)}\ \Big(o(n^{-1})+\ [(2n)!!/(l(n)(2n-1)!!)]\sum_{1\le k\le l(n)}\mathrm{O}\left(\sin^m((x-t_{k,l(n)})/2)\right)\\
&\quad\times\left(\cos^{2n}(x-t_{k,l(n)})/2\right)\sum_{0\le 2j\le[(m+2)/2]}P_{m+2,m+2-j}(n)\tan^{m+2-2j}(x-t_{k,l(n)})/2\Big)\\
&= h^{(m)}\ \Big(o(n^{-1})+\mathrm{O}\Big(\sum_{0\le 2j\le[(m+2)/2]}P_{m+2,m+2-j}(n)\\
&\qquad\qquad\times V_{n-\mu/2-1+j,l(n)}\left(\sin^{2m+2-2j}((x-t_{k,l(n)})/2);x\right)\Big)\Big)\\
&= h^{(m)}\ \Big(o(n^{-1})+\mathrm{O}\Big(\sum_{0\le 2j\le[(m+2)/2]}n^{m+2-j}(n^{-m-1+j})\Big)\Big)\\
&\le nL\ h^{(m)}\ ,
\end{aligned}
$$

where L is a constant. The result follows.

The following is the main direct and inverse result about approximation of f by $V_{n,l(n)}f$ (case $m = 0$) and its continuous derivatives $f^{(m)}$ by the m-th derivative $V_{n,l(n)}{}^{(m)}f$ of $V_{n,l(n)}f$:

Theorem 3.4.10.

Let $f(x)\in C_{2\pi}{}^{m}$, $m\ge 0$, $l(n) > n+[m/2]+2$, and $\psi\in$ S. Then:

(i) $V_{n,l(n)}{}^{(m)}f - f^{(m)} \le AK_{m,m+2}(f,n^{-1/2})\to 0, n\to\infty$, A being a constant;

(ii) $V_{n,l(n)}{}^{(m)}f - f^{(m)} = \mathrm{O}(\psi(n^{-1})), n\to\infty$, iff $K_{m,m+2}(f;t) = \mathrm{O}(\psi(t^2)), t\to 0$;

(iii) $V_{n,l(n)}{}^{(m)}f - f^{(m)} = o(\psi(n^{-1})), n\to\infty$, iff $K_{m,m+2}(f;t) = o(\psi(t^2)), t\to 0$;

(iv) $V_{n,l(n)}{}^{(m)}f - f^{(m)} \sim K_{m,m+2}(f;n^{-1/2})$, if $K_{m,m+2}(f;t^{1/2})\in S$.

Proof: (i)

$$
\begin{aligned}
V_{n,l(n)}{}^{(m)}f - f^{(m)} &\le V_{n,l(n)}{}^{(m)}(f-g) + V_{n,l(n)}{}^{(m)}g - g^{(m)} + g^{(m)} - f^{(m)}\\
&\le (1+K_m)\ f^{(m)} - g^{(m)}\ + n^{-1}M\ g^{(m+2)}\ .
\end{aligned}
$$

Taking infimum over $g\in C_{2\pi}{}^{m+2}$,

$$V_{n,l(n)}{}^{(m)}f - f^{(m)} \le AK_{m,m+2}(f;n^{-1/2}),$$

where $A = \max\{1+K, M\}$. To prove (ii)-(iv), using (i), if

$$K_{m,m+2}(f;t) = O(\psi(t^2)), t \to 0, \ V_{n,l(n)}{}^{(m)} f - f^{(m)} = O(\psi(n^{-1})), n \to \infty,$$

and, if

$$K_{m,m+2}(f;t) = o(\psi(t^2)), t \to 0, \ V_{n,l(n)}{}^{(m)} f - f^{(m)} = o(\psi(n^{-1})), n \to \infty.$$

For the inverse parts of (ii)- (iv), using the previous two Lemmas,

$$\begin{aligned} K_{m,m+2}(f;t) &\le K_{m,m+2}(f - V_{n,l(n)}(f;t)) + K_{m,m+2}(V_{n,l(n)}(f-g);t) + K_{m,m+2}(V_{n,l(n)}g;t) \\ &\le f^{(m)} - V_{n,l(n)}{}^{(m)} f + t^2 \left\{ V_{n,l(n)}{}^{(m+2)}(f-g) + V_{n,l(n)}{}^{(m+2)} g \right\} \\ &= f^{(m)} - V_{n,l(n)}{}^{(m)} f + t^2 \left\{ Ln\, f^{(m)} - g^{(m)} + K_{m+2}\, g^{(m+2)} \right\}. \end{aligned}$$

Taking infimum over $g \in C_{2\pi}{}^{m+2}$, for a constant M,

$$K_{m,m+2}(f;t) \le M \left\{ f^{(m)} - V_{n,l(n)}{}^{(m)} f + (tn^{1/2})^2 [K(f;n^{-1/2})] \right\},$$

and the results follow from Lemma 1.2.2.

The following consequence of a general interest is immediate:

Corollary 3.4.11.

Let $f(x) \in C_{2\pi}{}^{m}$, $m \ge 0$, $0 < \alpha < 2$, and $l(n) > n+[m/2]+2$. Then:

(i) $V_{n,l(n)}{}^{(m)} f - f^{(m)} = O(n^{-\alpha/2}), n \to \infty$, iff $K_{m,m+2}(f;t) = O(t^{\alpha}), t \to 0$;

(ii) $V_{n,l(n)}{}^{(m)} f - f^{(m)} = o(n^{-\alpha/2}), n \to \infty$, iff $K_{m,m+2}(f;t) = o(t^{\alpha}), t \to 0$;

(iii) $V_{n,l(n)}{}^{(m)} f - f^{(m)} \sim n^{-\alpha/2}$, if $K_{m,m+2}(f;t) \sim t^{\alpha}$.

Finally we relate the K-functional $K_{m,m+2}(f;t)$ and the second order modulus of continuity $\omega_2(f^{(m)};\delta)$ of the derivative $f^{(m)}(x)$, by showing that they are equivalent. (The case $m = 0$ of the result is, of course, well known. However our proof is valid for all $m \ge 0$).

Theorem 3.4.12.

For $m \ge 0$, and $f \in C_{2\pi}{}^{m}$,

$$\omega_2(f^{(m)};t) \le 4K_{m,m+2}(f;t) \le 9\omega_2(f^{(m)};t), t > 0.$$

Proof: Let $g \in C_{2\pi}^{m+2}$. Then,

$$\Delta_h^2 f^{(m)}(x) = \Delta_h^2 \{f^{(m)}(x) - g^{(m)}(x)\} + \int_0^h \int_0^h g^{(m+2)}(x+r+s)\,dr\,ds.$$

Hence,

$$\omega_2(f^{(m)};h) \le 4\,\|f^{(m)} - g^{(m)}\| + h^2\,\|g^{(m+2)}\|.$$

Taking infimum over $g \in C_{2\pi}^{m+2}$, we get

$$\omega_2(f^{(m)};t) \le 4K_{m,m+2}(f;t).$$

Conversely, with δ_h denoting the central difference operator with step h, and choosing

$$g(x) = (2h)^{-2} \int_{-h}^{h} \int_{-h}^{h} f(x+r+s)\,dr\,ds,$$

$$\begin{aligned} g^{(m+2)}(x) &= \left[(2h)^{-2} \int_{-h}^{h} \int_{-h}^{h} f^{(m)}(x+r+s)\,dr\,ds\right]'' \\ &= \left[(2h)^{-2} \int_{-h}^{h} \delta_{2h} f^{(m)}(x+s)\,ds\right]' \\ &= (2h)^{-2} \delta_{2h}^2 f^{(m)}(x). \end{aligned}$$

It follows that

$$\|g^{(m+2)}\| \le (2h)^{-2} \omega_2(f^{(m)};2h),$$

and taking $h = t/2$, we have

$$\|g^{(m+2)}\| \le (t)^{-2} \omega_2(f^{(m)};t)$$

Also, we have

$$\begin{aligned} f^{(m)}(x) - g^{(m)}(x) &= f^{(m)}(x) - \left[(2h)^{-2} \int_{-h}^{h} \int_{-h}^{h} f^{(m)}(x+r+s)\,dr\,ds\right] \\ &= f^{(m)}(x) - \left[(2h)^{-2} \int_0^h \int_0^h \{f^{(m)}(x+r+s) + f^{(m)}(x-r+s) \right. \\ &\qquad \left. + f^{(m)}(x+r-s) + f^{(m)}(x-r-s)\}\,dr\,ds\right] \\ &= -(2h)^{-2} \int_0^h \int_0^h \{\delta_{r+s}^2 f^{(m)}(x) + \delta_{|r-s|}^2 f^{(m)}(x)\}\,dr\,ds, \end{aligned}$$

from which it follows that

$$4\,\|f^{(m)} - g^{(m)}\| \le \omega_2(f^{(m)};2h) + \omega_2(f^{(m)};h) \le 5\omega_2(f^{(m)};h).$$

Hence,

$$4K_{m,m+2}(f;t) \le 4\{\,\|f^{(m)} - g^{(m)}\| + t^2\,\|g^{(m+2)}\|\,\} \le 9\omega_2(f^{(m)};h),$$

completing the proof.

As an immediate consequence, the earlier direct and inverse results for the discrete De La Vallée-Poussin filters could be alternately re-stated as:

Corollary 3.4.13.

Let $f(x) \in C_{2\pi}^{m}$, $m \geq 0$, $l(n) > n+[m/2]+2$, $\psi \in S$, and $0 < \alpha < 2$. Then:

(i) $V_{n,l(n)}{}^{(m)} f - f^{(m)} \leq A\omega_2(f, n^{-1/2}) \to 0, n \to \infty$, A being a constant;

(ii) $V_{n,l(n)}{}^{(m)} f - f^{(m)} = O(\psi(n^{-1})), n \to \infty$, iff ;

(iii) $V_{n,l(n)}{}^{(m)} f - f^{(m)} = o(\psi(n^{-1})), n \to \infty$, iff $\omega_2(f;t) = o(\psi(t^2)), t \to 0$;

(iv) $V_{n,l(n)}{}^{(m)} f - f^{(m)} = O(n^{-\alpha/2})$, iff $\omega_2(f;t) = O(t^{\alpha})$;

(v) $V_{n,l(n)}{}^{(m)} f - f^{(m)} = o(n^{-\alpha/2})$, iff $\omega_2(f;t) = o(t^{\alpha})$;

(vi) $V_{n,l(n)}{}^{(m)} f - f^{(m)} \sim n^{-\alpha/2}$, if $\omega_2(f;t) \sim t^{\alpha}$.

3.5. NUMERICAL SIMULATIONS AND DISCUSSION FOR KOROVKIN, JACKSON AND GENERALIZED JACKSON OPERATORS

Numerical simulation results for Korovkin, Jackson and Generalized Jackson operators with n = 128, 64, 32, 16 and 8 are presented in this section. The experiments were done on test image set I.

In Figure 3.1 (Original T_2 weighted and corresponding images approximated with Korovkin operator), the first image is original T_2 weighted image and other images correspond to the Korovkin operator with n = 128, 64, 32, 16 and 8 respectively. Figure 3.2 (MLE Segmented images of original and corresponding images approximated with Korovkin operator) contains MLE-based segmented images associated with the images in Figure 3.1. The number of iterations taken for original image is 59400 and for n = 128, 64, 32, 16, 8 are 10250, 36150, 2950, 2750, 3000 respectively.

In Figure 3.3 (Original T_2 weighted and corresponding images approximated with Jackson operator), the first image is original T_2 weighted image and other images correspond to the Jackson operator with n = 128, 64, 32, 16 and 8 respectively. Figure 3.4. (MLE Segmented images of original and corresponding images approximated with Jackson operator) contains MLE-based segmented images associated with the images in Figure 3.3. The number of iterations taken for original image is 59400 and for n = 128, 64, 32, 16, 8 are 14400, 10200, 25900, 9000, 2550 respectively.

In Figure 3.5 (Original T_2 weighted and corresponding images approximated with Generalized Jackson operator), the first image is original T_2 weighted image and other images correspond to the Generalized Jackson operator with n = 128, 64, 32, 16 and 8 respectively. Figure 3.6. (MLE Segmented images of original and corresponding images approximated with Generalized Jackson operator) contains MLE-based segmented images associated with the images in Figure 3.5. The number of iterations taken for original image is 59400 and for n = 128, 64, 32, 16 and 8 are 8450, 16150, 22150, 9200 and 2900 respectively.

of iterations taken for original image is 59400 and for n = 128, 64, 32, 16, 8 are 8450, 16150, 22150, 9200, 2900 respectively.

The percentage relative errors in the approximation with Korovkin, Jackson and Generalized Jackson operators are given in Tables 3.1 and 3.2.

n	Korovkin	Jackson	Generalized Jackson
8	18.385885	18.120655	17.940020
16	18.066454	17.019840	16.342531
32	15.425109	13.358624	11.905014
64	11.264680	8.087324	6.817662
128	4.510457	3.271766	1.597230

Table 3.1: % Relative L1 Errors in the Approximation of the Original Cross-Section with Korovkin, Jackson and Generalized Jackson Operators

n	Korovkin	Jackson	Generalized Jackson
8	20.188421	19.894699	19.693750
16	19.834602	18.656040	17.878233
32	16.788383	14.335914	12.556752
64	11.735953	7.875902	6.418863
128	4.802420	2.915334	1.400528

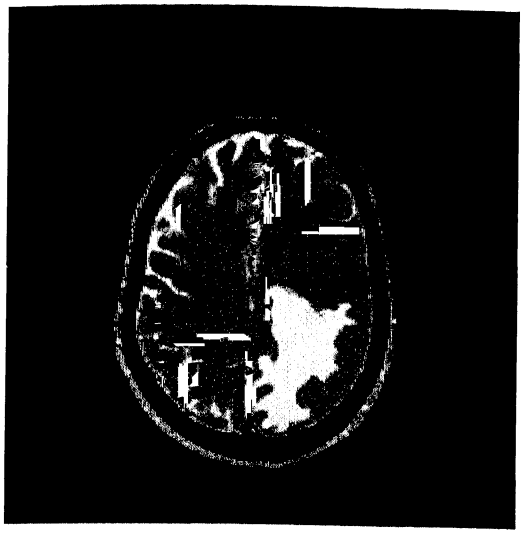
Table 3.2: % Relative L2 Errors in the Approximation of the Original Cross-Section with Korovkin, Jackson and Generalized Jackson Operators

The case n = 128 gives an image visually similar to the original and the case n = 64 gives necessary smoothing without distorting major regions with all operator mentioned above. Blurring in the images increases as n decreases from 128 to 8 and it is high for n = 32 downwards. The blurring with korovkin operator is more followed by Jackson operator and then Generalized Jackson operator. This is clear from Tables 3.1 and 3.2 and is also observed from the images shown in Figures 3.1, 3.3 and 3.5.

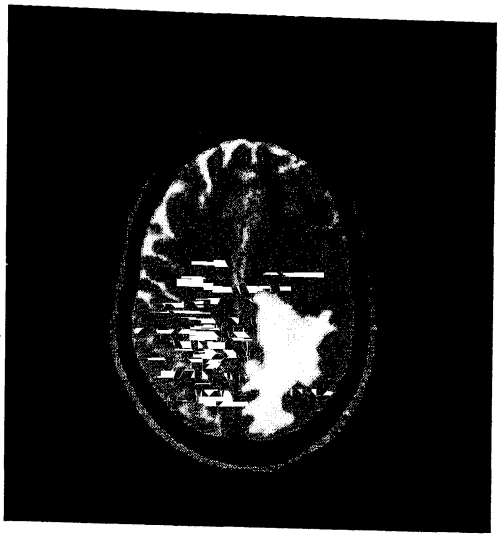
The segmented image for n = 128 is similar to that obtained from original cross-section, result for n = 64 seems to be satisfactory segmentation of major tissues

along with the pathology. For n = 32, 16 and 8, the boundary between major segments like cranial part of the brain, white matter and edema, thickens and smaller regions merge with neighboring larger segments. This is observed that the segmented image obtained for n = 16 (Figure 3.2) with korovkin operator matches with that obtained for n = 8 with Jackson (Figure 3.4) and Generalized Jackson (Figure 3.6) operators.

With korovkin operator, segmented image obtained for n = 64, a thin boundary encloses the region comprising of edema completely which could is useful for medical purposes. Also, for n = 64, 32 and 16, the segmented images show three regions corresponding to edema which are represented as different growth stages of an edema within the brain parenchyma. Similar kind of splitting of edema has also been observed in the cases of Jackson and Generalized Jackson operators for n = 32, 16 and 8. In cases of these two operators, boundary enclosing the edema is observed for n = 32. Plots of percentage relative errors in the approximation with respect to the values of n are given in Graph 3.1.

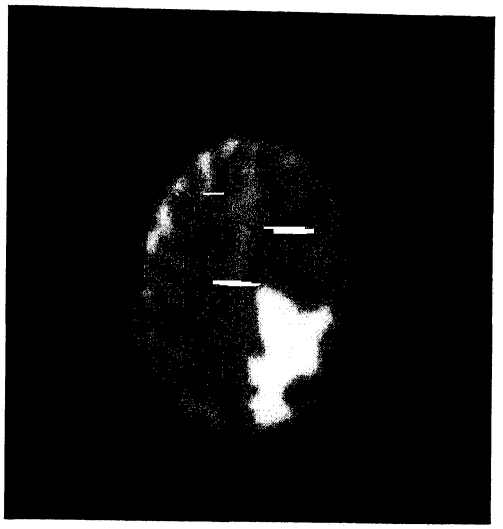


Original



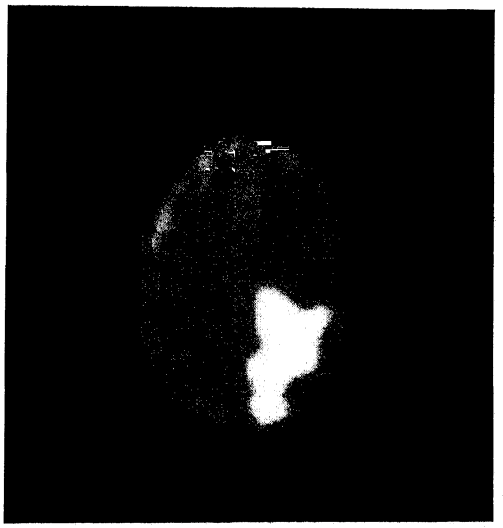
n=128

n=64



n=32

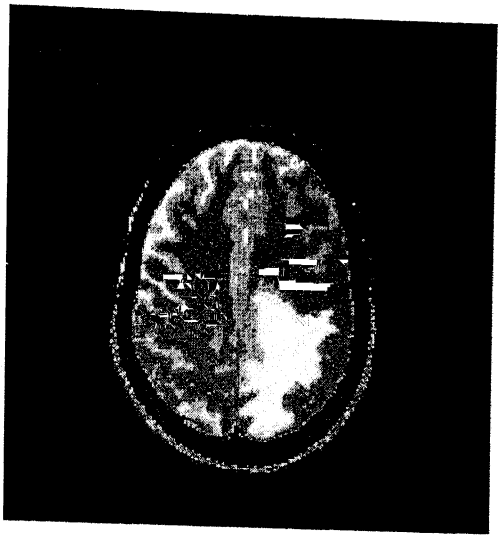
n=16



n=8

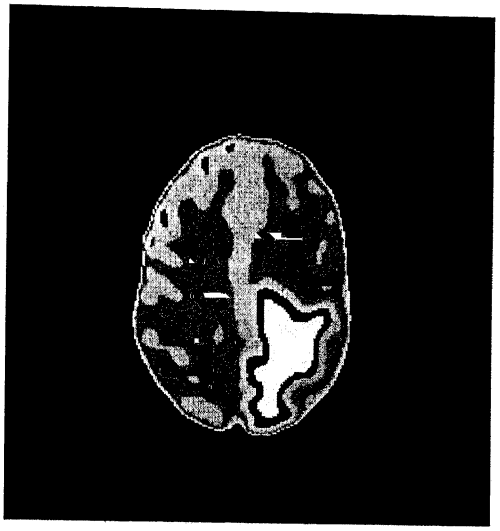
Figure 3.1: Original T_2 weighted and Corresponding Images Approximated with Korovkin Operator

Original

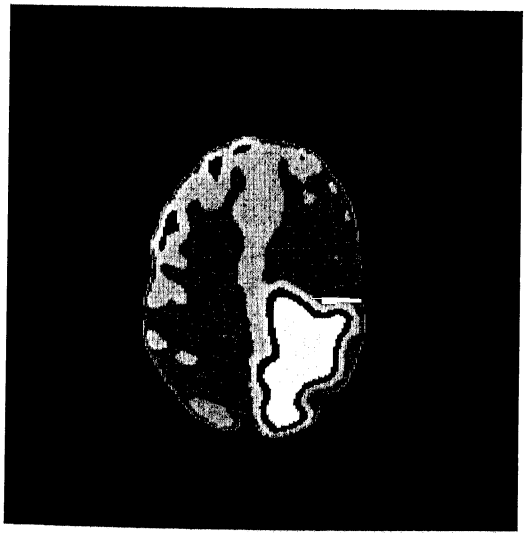


n=128

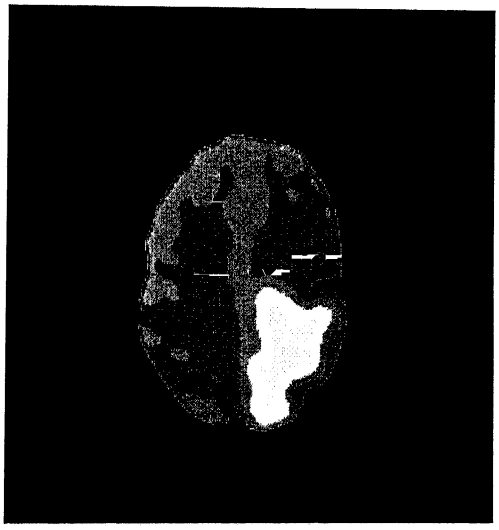
n=64



n=32

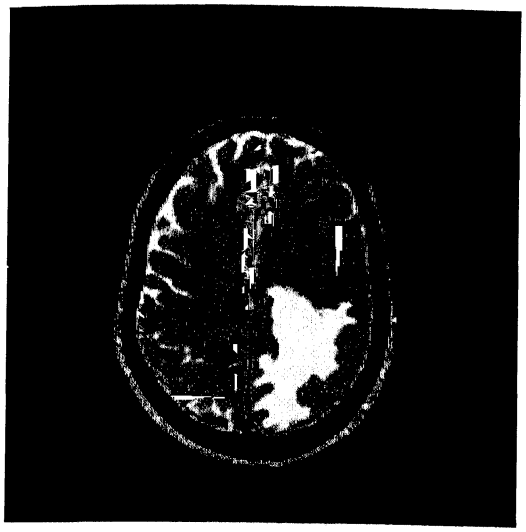


n=16

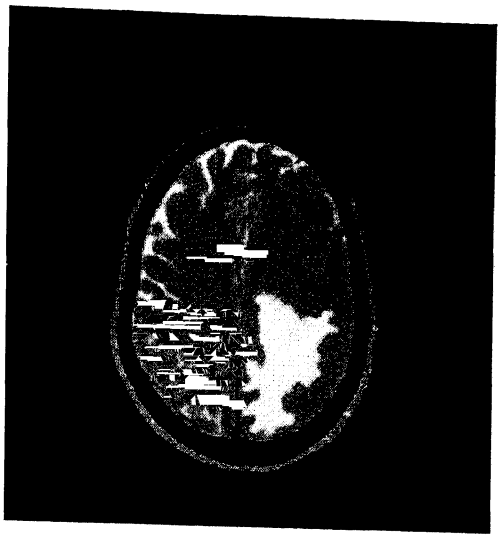


n=8

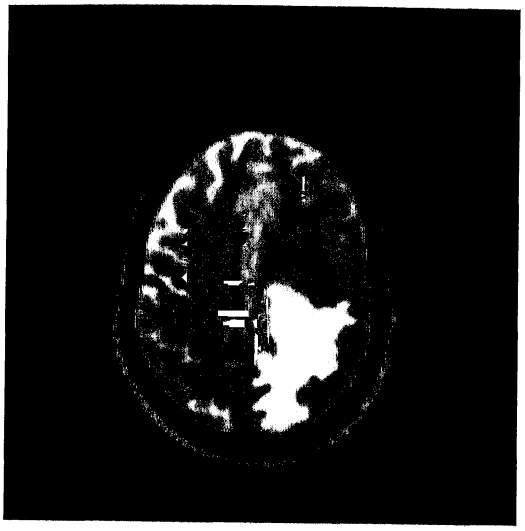
Figure 3.2: MLE Segmented Images of Original and Corresponding Images in Figure 3.1



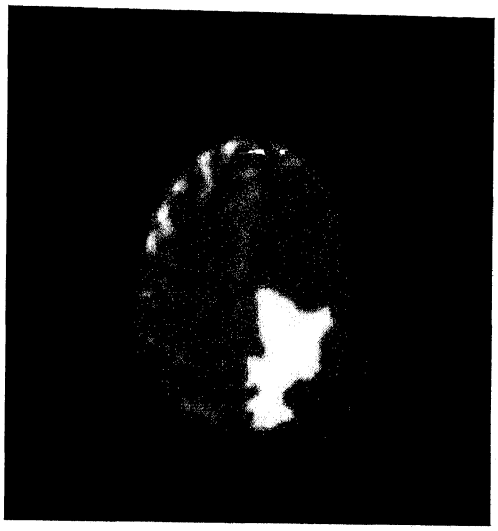
Original



n=128

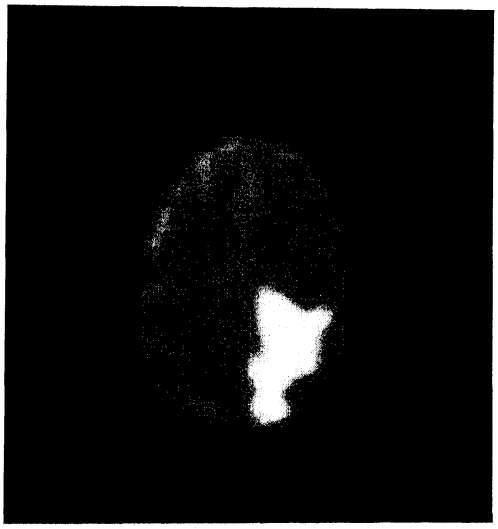


n=64



n=32

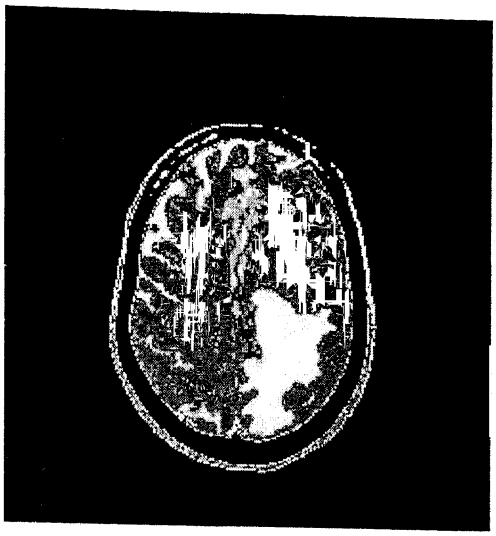
n=16



n=8

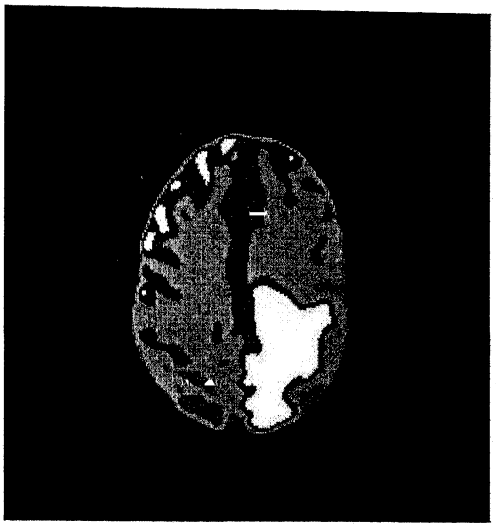
Figure 3.3: Original T_2 weighted and Corresponding Images Approximated with Jackson Operator

Original



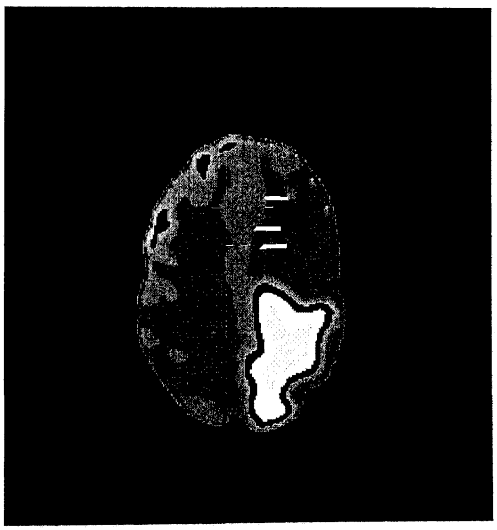
n=128

n=64



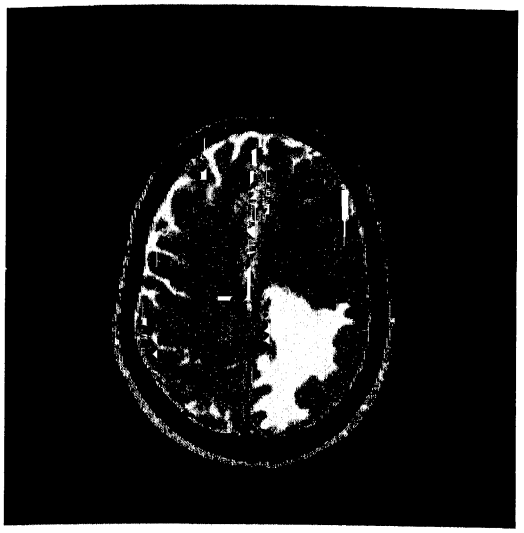
n=32

n=16

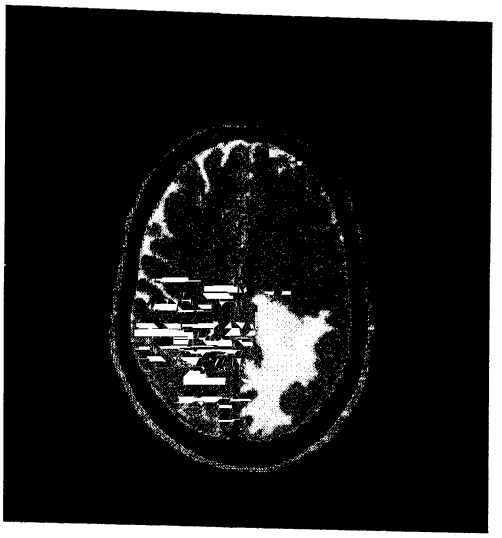


n=8

Figure 3.4: MLE Segmented Images of Original and Corresponding Images in Figure 3.3

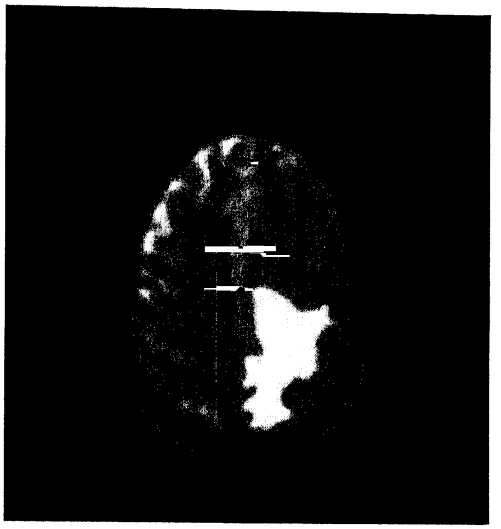


Original

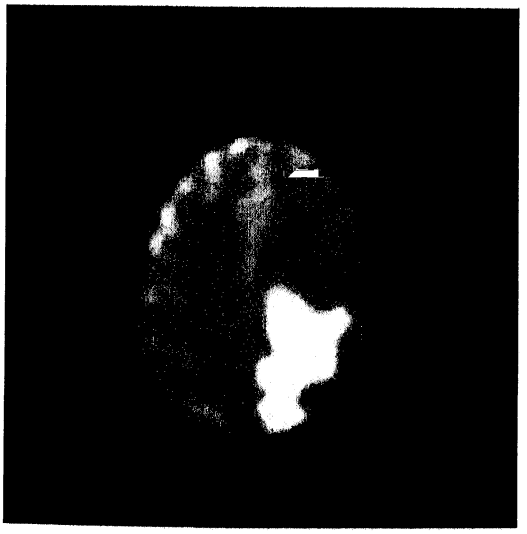


n=128

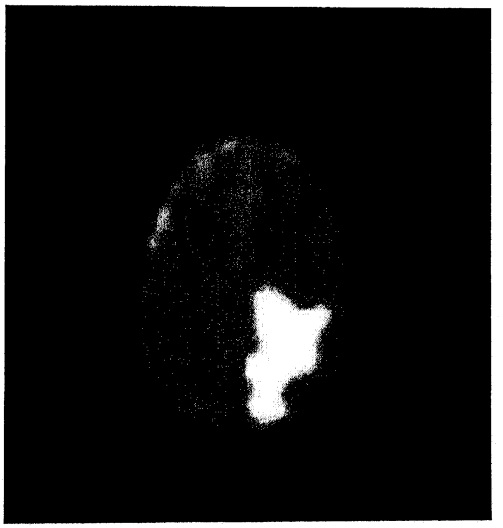
n=64



n=32



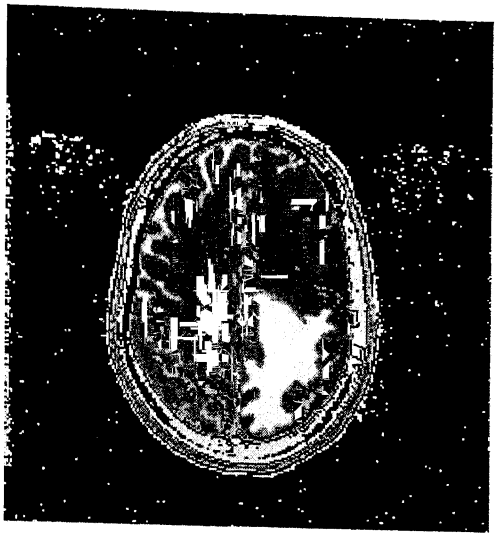
n=16



n=8

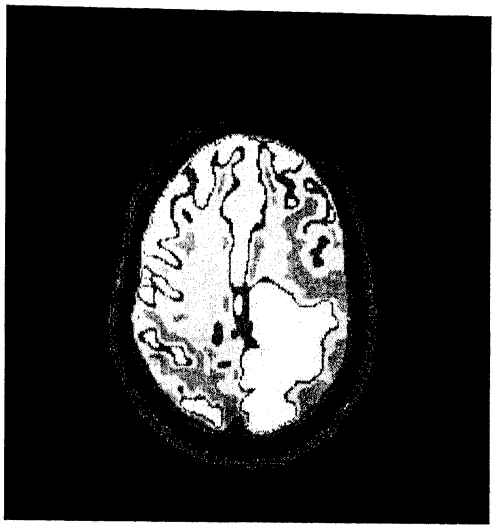
Figure 3.5: Original T_2 weighted and Corresponding Images Approximated with Generalized Jackson Operator

Original

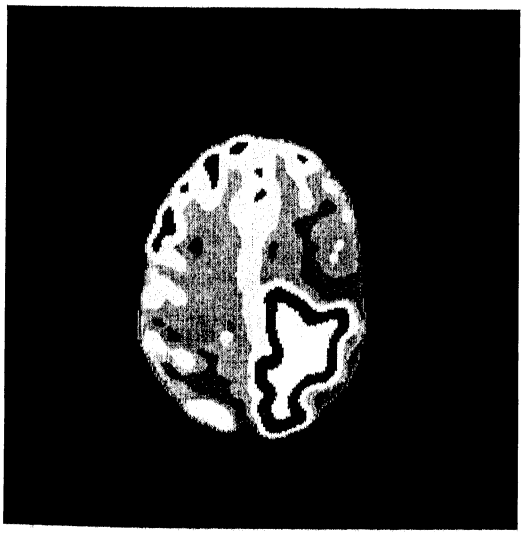


n=128

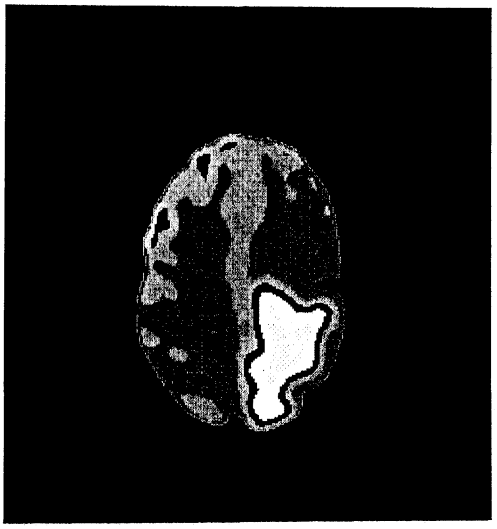
n=64



n=32

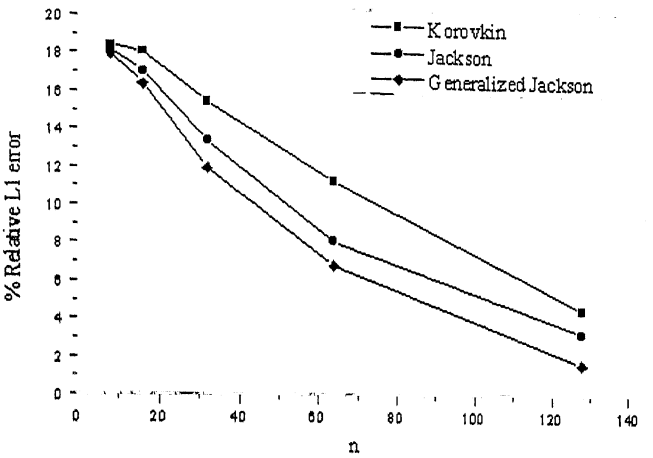


n=16

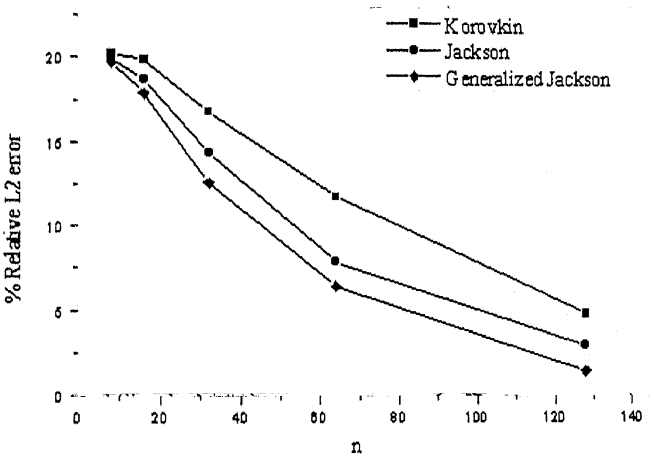


n=8

Figure 3.6: MLE Segmented Images of Original and Corresponding Images in Figure 3.5



% Relative L1 Error Vs n



% Relative L2 Error Vs n

Graph 3.1: Plots of % Relative L1 and L2 Errors with respect to the Values of n

3.6. NUMERICAL SIMULATIONS AND DISCUSSION CORRESPONDING TO VALLÉE-POUSSIN KERNELS

Numerical simulation results for De La Vallée-Poussin operator with n = 4096, 2048, 1024, 512 and 256 are presented in this section. The experiments were done on test image set I.

In Figure 3.7 (Original T_2 weighted and corresponding images approximated with De La Vallée-Poussin operator), the first image is original T_2 weighted image and other images correspond to the De La Vallée-Poussin operator with n = 4096, 2048, 1024, 512 and 256 respectively. Figure 3.8 (MLE Segmented images of original and corresponding images approximated with De La Vallée-Poussin operator) contains MLE-based segmented images associated with the images in Fig. 3.7. The number of iterations taken for original image is 59400 and for n = 4096, 2048, 1024, 512, 256 are 10850, 11850, 7500, 9450, 5400 respectively.

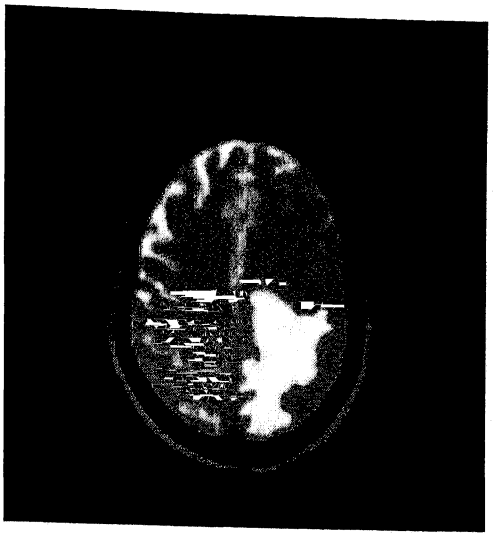
The percentage L1 and L2 percentage errors in the approximation for De La Vallée-Poussin Kernel are summarized in Table 3.3.

n	4096	2048	1024	512	256
% Relative L1 Error	9.52296	12.16614	14.63782	16.38696	17.40005
% Relative L2 Error	9.62524	12.87752	15.87117	17.92870	19.08679

Table 3.3: % Relative L1 and L2 Errors in the Approximation of the Original Cross-Section with De La Vallée-Poussin Kernel

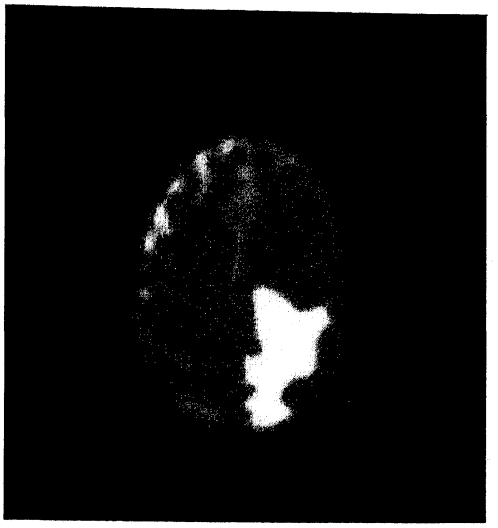
In case of De La Vallée-Poussin operator for n = 128, 64, 32, 16 and 8, the smoothing increases which leads to a segmentation of interior brain region and its outer cranial part as two major segments. Interior part of the brain mainly represents two regions corresponding white matter and a combination of CSF and gray matter. With this operator also, the region represented by edema splits into two and then three parts as smoothing increases from n = 4096 to n = 256 and the boundary surrounding the edema thickens as smoothing increases.

Original



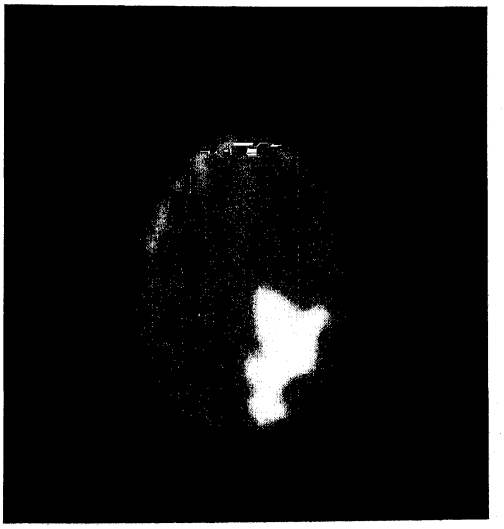
n=4096

n=2048



n=1024

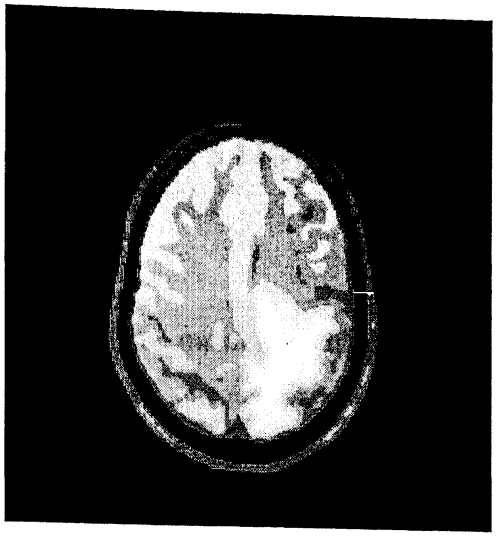
n=512



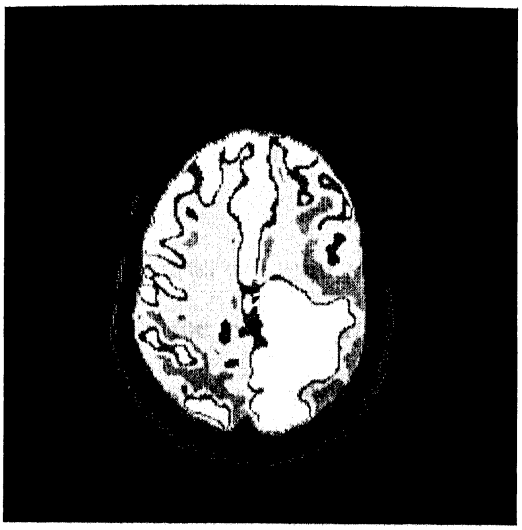
n=256

Figure 3.7: Original T_2 weighted and Corresponding Images Approximated with De La Vallée-Poussin Kernel

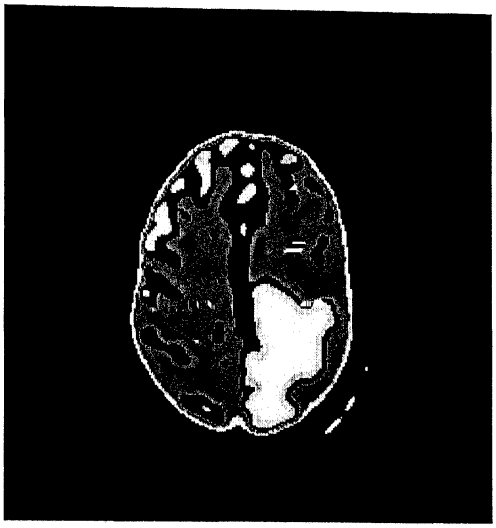
Original



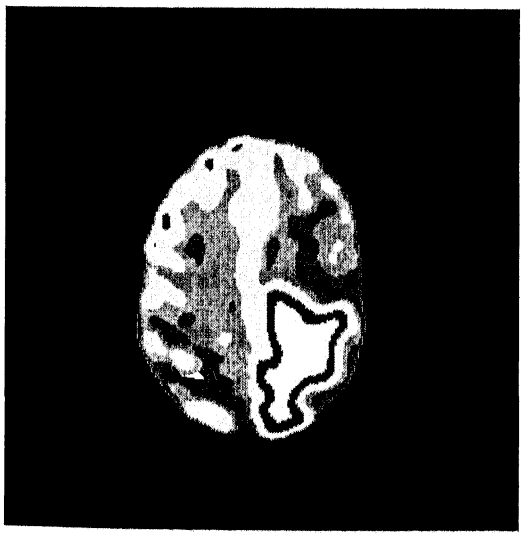
n=4096



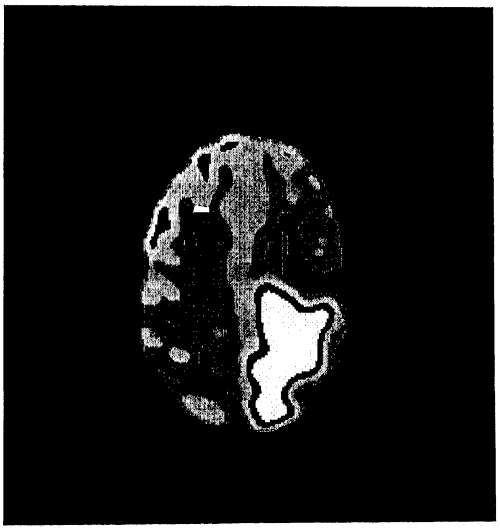
n=2048



n=1024



n=512



n=256

Figure 3.8: MLE Segmented Images of Original and Corresponding Images in Figure 3.7

CHAPTER IV : SEGMENTATION METHODS BASED ON MAXIMUM LIKELIHOOD ESTIMATION

In this chapter, segmentation methods based on maximum likelihood estimation (MLE) have been studied. These methods have already been used in Chapters I, II and III in the numerical simulations. We have observed elsewhere that in MR images the distributions of pixel intensities of CSF, and normal tissues like white and gray matter, as well as different pathologies strongly seem to follow Gaussian distributions (R.K.S. Rathore, S. Datta, R.K. Gupta, S.B. Rao and R. Verma [2001]).

Hence in the following treatment, the problem of segmentation has been modeled by considering the pixel intensities as random variables and the likelihood function defined as the joint probability density $f(x, \theta, \mu, \sigma)$ of these variables, which is a function of the vector parameters $\theta = (\theta_1, \theta_2, \ldots, \theta_m)'$, $\mu = (\mu_1, \mu_2, \ldots, \mu_m)'$ and $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_m)'$. Considering the normal distribution of each tissue class within a pixel, the density function of the pixel intensity x is given as:

$$f(x;\theta,\mu,\sigma) = \sum_{j=1}^{m} \frac{\theta_j}{\sigma_j\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu_j)^2}{2\sigma_j^2}\right\}.$$

4.1. INTRODUCTION

The general method of maximum likelihood estimation is as follows: Let X denote the realized value of a set of observations and P(X,θ) denote the joint density, where $\theta = (\theta_1, \ldots, \theta_q)$, the vector of parameters belongs to a set $\Theta \subset R_q$, where R_q is the q-dimensional real space. The likelihood of θ given the observations is defined to be a function of θ:

$$L(\theta|X) \propto P(X, \theta).$$

The principle of maximum likelihood consists of accepting $\theta^* = (\theta_1^*, \ldots, \theta_q^*)$ as the estimate of θ, where

$$L(\theta^*|X) = \sup_{\theta \in \Theta} L(\theta|X).$$

In practice it is convenient to work with $l(\theta|X) = \text{Log } L(\theta|X)$, in which case θ^* satisfies the equation

$$l(\theta^*|X) = \sup_{\theta \in \Theta} l(\theta|X).$$

If the supremum is attained at an interior point of Θ and $l(\theta|X)$ is a differentiable function of θ, then the partial derivatives vanish at that point, so that θ^* is a solution of the equations

$$\frac{\partial l(\theta \mid X)}{\partial \theta_i} = 0, \qquad i = 1, \ldots, q.$$

These equations are called *maximum likelihood equations* and any solution of them as a *maximum likelihood estimate* (C.R. Rao [1989]).

Suppose M pixels on an image are distributed into m segments with θ_j as the jth population size ratio corresponding to the tissue class j (so that $\Sigma_{1 \le j \le m} \theta_j = 1$). Then, the likelihood function for the totality of the pixel intensities, modeled as the mixture of the m number of normal populations is given by,

$$f(x;\theta,\mu,\sigma) = \prod_{i=1}^{M} \left\{ \sum_{j=1}^{m} \theta_j N(x_i;\mu_j,\sigma_j) \right\},$$

where μ_j is the mean, σ_j the standard deviation, and θ_j is the size ratio of the jth population, x denotes the pixel intensity and the Gaussian density function is given by

$$N(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right].$$

This single image case of the present model is discussed in the next section 4.2, while the generalization of the same to the multispectral case of a set of images for the same cross-section is considered in section 4.3. The numerical and clinical details of the corresponding results in the test cases are presented in the later sections of the chapter.

4.2. SINGLE CONTRAST SEGMENTATION METHOD

In this section, we consider a single contrast normal distribution as follows:

$$L = \prod_{i=1}^{M} \left(\sum_{j=1}^{m} \theta_j N(x_i; \mu_j, \sigma_j) \right).$$

The objective function $\log L$ is maximized with respect to $\sum_{j=1}^{m} \theta_j = 1$. In this case, the maximum likelihood function is obtained via calculating the estimators for which the likelihood function is maximum by solving the normal equations $\partial L/\partial \mu_k = 0$, $\partial L/\partial \sigma_k = 0$ and $\partial L/\partial \theta_k = 0$, for $k = 1, \ldots, m$, or by solving the normal equations $\partial \log L/\partial \mu_k = 0$, $\partial \log L/\partial \sigma_k = 0$ and $\partial \log L/\partial \theta_k = 0$, for $k = 1, \ldots, m$, as both $L(\theta)$ and $\log L(\theta)$ have maxima at the same values of μ_k, σ_k and θ_k, $k = 1, \ldots, m$. Here μ_k, σ_k and θ_k represent the corresponding mean, standard deviation and the population size ratio of the kth segment present on an image. Let M denote the total number of pixels on an image. The following numerical iterative scheme (R.K.S. Rathore, S. Datta, R.K. Gupta, S.B. Rao and R. Verma [2001]) is easily seen to be consistent with the maximum likelihood estimation (MLE) normal equations:

$$\mu_k^{(n+1)} = (1/D_n) \sum_{i=1}^{M} \theta_k^{(n)} N(x_i; \mu_k^{(n)}, \sigma_k^{(n)}) x_i / S_{i,n}, \tag{4.2.1}$$

$$\sigma_k^{(n+1)} = \sqrt{(1/D_n) \sum_{i=1}^{M} \theta_k^{(n)} N(x_i; \mu_k^{(n)}, \sigma_k^{(n)}) (x_i - \mu_k^{(n)})^2 / S_{i,n}}, \tag{4.2.2}$$

$$\theta_k^{(n+1)} = (1/M) \sum_{i=1}^{M} \theta_k^{(n)} N(x_i; \mu_k^{(n)}, \sigma_k^{(n)}) / S_{i,n}, \tag{4.2.3}$$

where

$$D_n = \sum_{i=1}^{M} \theta_k^{(n)} N(x_i; \mu_k^{(n)}, \sigma_k^{(n)}) / S_{i,n}, \tag{4.2.4}$$

and

$$S_{i,n} = \sum_{j=1}^{m} \theta_j^{(n)} N(x_i; \mu_j^{(n)}, \sigma_j^{(n)}). \tag{4.2.5}$$

Iterative computation of means, standard deviations and population size ratios and pixel classification in the single image case, using above equations (4.2.1-5) could be summarized as the following algorithm:

Algorithm 4.2.1.

Step 1: Given that m segments are present on an image, start with some initial estimates $\mu_k^{(0)}$, $\sigma_k^{(0)}$ and $\theta_k^{(0)}$, $k = 1, \ldots, m$, of their respective means, standard deviations and the population size ratios.

Step 2: Update means $\mu_k^{(n)}$, standard deviations $\sigma_k^{(n)}$ and population size ratios $\theta_k^{(n)}$ using equations 4.2.1-4.2.5, till $\theta_k^{(n+1)} - \theta_k^{(n)} < A\theta_k^{(n)}$, $k = 1, \ldots, m$, for some tolerance constant $A > 0$. (Note that k stands for a segment number and n for an iteration count.)

Step 3: Calculate the probability function using the normal density function with respect to each of the segments for each pixel and classify the pixel to that segment for which the probability function is maximum.

4.3. MULTI-RESOLUTION SEGMENTATION METHOD

The contrast between the normal tissues like white matter, gray matter and the cerebrospinal fluid (CSF) is more on T_2 and T_1 weighted images as compared to that on the proton density (PD) weighted images. T_2 weighted images show better contrast between white and gray matter, while T_1 weighted images give a better contrast of CSF with that of the brain parenchyma (white matter and gray matter). Experience shows that it is not possible to get satisfactory segmentations using simple thresholding algorithms or similar such algorithms based on intensity only. This is because, in addition to the images not having a consistently good intra-tissue contrast across the entire boundaries, the presence of inhomogeneities within different tissue types results in the overlaps of different regions at their boundaries, as well as in the appearance of perforations in their interiors. The variability of the contrast in different tissue-types obtained by employing different sequences of magnetic resonance images suggests using multivariate analysis for the segmentation of a region of interest in a cross-section.

Consider pixel intensity as a multivariate possessing different correlations between the pairs of given weighted images. The above methodology is not applicable in the present case and therefore the following iterative scheme is applicable to only cases comprising of two or more images.

The derivation of the iterative scheme in this section is based on a multivariate normal distribution modeling of the mixutre of the tissue populations. From the definition of likelihood function, the model could be defined as:

$$L = \prod_{i=1}^{M} \left(\sum_{j=1}^{m} \theta_j N(\mathbf{x}_i; \mu_j, \Sigma_j) \right),$$

where

$$N(\mathbf{x}_i; \mu_j, \Sigma_j) = \frac{1}{(2\pi)^{n/2} |\Sigma_j|^{1/2}} \exp\left\{ -\frac{1}{2} (\mathbf{x}_i - \mu_j)' \Sigma_j^{-1} (\mathbf{x}_i - \mu_j) \right\}.$$

The main aim is to estimate the mean vectors, covariance matrices and the population size ratio vector $\theta = (\theta_1, \theta_2, \ldots, \theta_m)$ which maximize log L subject to the condition $\sum_{j=1}^{m} \theta_j = 1$. This method is applicable to n number of images, where n is greater than one and is finite. In this model, $\mathbf{x}_i$ is a $n \times 1$ vector for each pixel ($i = 1, \ldots, M$) on a cross-section, μ_i is an $n \times 1$ vector for each $j = 1, \ldots, m$, where m is the pre-assumed number of segments present on the cross-section and M is total number of pixels on it. Σ_j is an $n \times n$ positive semi-definite covariance matrix for each $j = 1, \ldots, m$. We have considered the covariance matrices to be strictly positive definite. In this case also, the maximum likelihood estimators are obtained by solving the normal equations so obtained by using partial differentiations with respect to the estimators.

The iteratively estimated mean vector $\mu_k^{(v+1)}$ of multivariate distribution may be obtained by solving system of linear equations given by

$$\mathbf{A}^{(v)} \mu_k^{(v+1)} = \mathbf{B}^{(v)}, \tag{4.3.1}$$

where the elements of $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$ for $r, s = 1, 2, \ldots, n$ are as follows:

$$a_{rs}^{(v)} = \sum_{i=1}^{M} \left(P_{ik}^{(v)} S_{krs}^{(v)} \Big/ \sum_{j=1}^{m} P_{ij}^{(v)} \right), \quad b_s^{(v)} = \sum_{i=1}^{M} \left(P_{ik}^{(v)} \sum_{p=1}^{n} S_{kps}^{(v)} x_{ip} \Big/ \sum_{j=1}^{m} P_{ij}^{(v)} \right)$$

where $P_{ij}^{(v)} = \theta_j^{(v)} N(\mathbf{x}_i; \mu_j^{(v)}, \Sigma_j^{(v)})$, and $S_{krs}^{(v)}$ are the elements of matrix $\Sigma_k^{-1(v)}$. For $k = 1, 2, \ldots, m$, the values of $\theta_k^{(v+1)}$ are obtained as

$$\theta_k^{(v+1)} = (1/M)\sum_{i=1}^{M} \left(P_{ik}^{(v)} \Big/ \sum_{j=1}^{m} P_{ij}^{(v)} \right). \tag{4.3.2}$$

Due to positive definiteness of covariance matrix $\left(\sum_j^{(v)}\right)^{-1}$, it can be decomposed into the product of a lower triangular matrix with its transpose using Cholesky decomposition. To compute the elements of covariance matrix $\left(\sum_j^{(v)}\right)^{-1}$, we first estimated the elements of the lower triangular matrix of its Cholesky decomposition. The ratios of estimated off-diagonal elements to the estimated diagonal elements of the lower triangular matrix are obtained from the n-1 systems of linear equations given by

$$\mathbf{C}_q^{(v)} \mathbf{z}_q^{(v+1)} = \mathbf{D}_q^{(v)}, q = 1, \ldots, n-1, \tag{4.3.3}$$

where, for $r, p > q$, the elements of $\mathbf{C}_q^{(v)}$ and $\mathbf{D}_q^{(v)}$ are computed as follows:

$$c_{rp}^{(v)} = \sum_{i=1}^{M} \left(P_{ik}^{(v)} (x_{ir} - \mu_{kr}^{(v)})(x_{ip} - \mu_{kp}^{(v)}) \Big/ \sum_{j=1}^{m} P_{ij}^{(v)} \right)$$

and

$$d_p^{(v)} = -\sum_{i=1}^{M} \left(P_{ik}^{(v)} (x_{iq} - \mu_{kq}^{(v)})(x_{ip} - \mu_{kp}^{(v)}) \Big/ \sum_{j=1}^{m} P_{ij}^{(v)} \right).$$

The estimated diagonal elements for q = 1, 2, ..., n of the lower triangular matrix are obtained as

$$c_{qq}^{(v)} = \sqrt{\sum_{i=1}^{M} \left(P_{ik}^{(v)} \left\{ \sum_{r>q} z_r^{(v)} (x_{ir} - \mu_{kr}^{(v)})(x_{iq} - \mu_{kq}^{(v)}) + (x_{iq} - \mu_{kq}^{(v)})^2 \right\} \Big/ \sum_{j=1}^{m} P_{ij}^{(v)} \right) \Big/ \sum_{i=1}^{M} \left(P_{ik}^{(v)} \Big/ \sum_{j=1}^{m} P_{ij}^{(v)} \right)} \tag{4.3.4}$$

Computation of mean vectors, covariance matrices and the population size ratios and pixel classification in the multi-resolution case could be summarized as the following algorithm:

Algorithm 4.3.1.

Step 1: Given that m segments are present on an image, start with some initial estimates of their respective mean vectors, covariance matrices and the population size ratios.

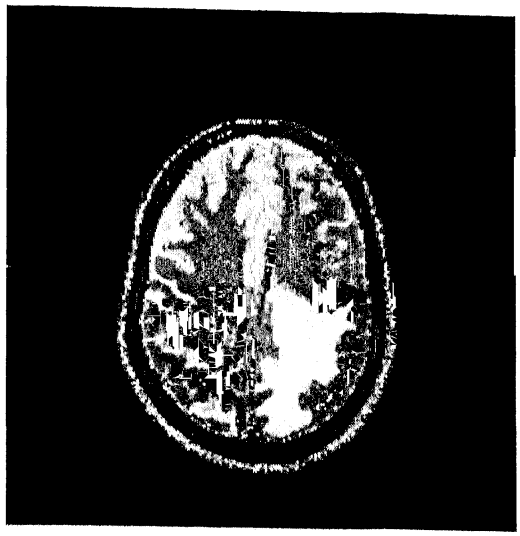
Step 2: Estimate mean vectors, covariance matrices, and the population size ratios using equations (4.3.1-4.3.4), till $\theta_k^{(v+1)} - \theta_k^{(v)} < B\theta_k^{(v)}, k = 1,\ldots m$, for some tolerance constant $B > 0$. (Here k denotes the segment, and v the iteration number.)

Step 3: Calculate the probability function using the multivariate normal density function with respect to each of the segments for each pixel and classify the pixel to that segment for which the probability function is maximum.

4.4. OBSERVATIONS RELATED TO TEST IMAGE SETS I AND II

The segmented cross-section obtained using Algorithm 4.2.1 on test image set I with A = 0.001 is shown in Figure 4.1. The segmented cross-section obtained using Algorithm 4.2.1 on T_2-weighted image of set II with A = 0.001 is given in Figure 4.2, in which ventricles and edema both fall under a single distribution.

The segmented images so obtained on application of Algorithm 4.3.1. to all the possible pairs of images of image set II, with B = 0.02, are shown in Figure 4.3. Ventricles and edema merge to give a single segment in segmented image obtained using PD and T_2 weighted images, while proper segmentation of white matter and a part of edema adjacent to ventricles are found in segmented image obtained using PD and T_1 weighted images, which also shows two different segments in ventricles. Only one segment represents both edema and ventricles in segmented image obtained using T_2 and T_1 weighted images, which also shows a part of white matter merging with gray matter.



T_2 weighted Image

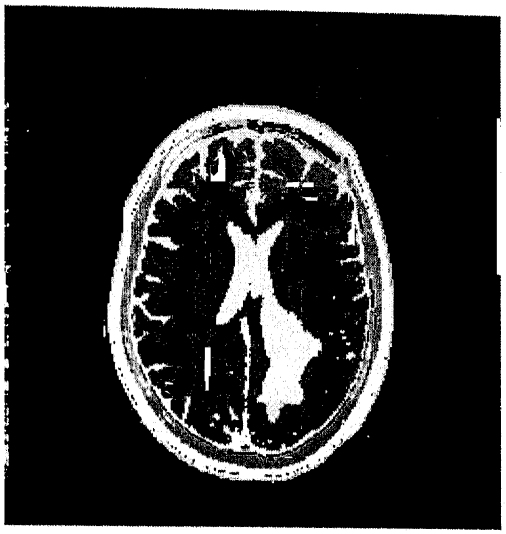
Figure 4.1: Segmented Image Obtained using Algorithm 4.2.1 to Test Image Set I

T_2 weighted Image

Figure 4.2: Segmented Image Obtained using Algorithm 4.2.1 to T_2 weighted Image of Test Image Set II

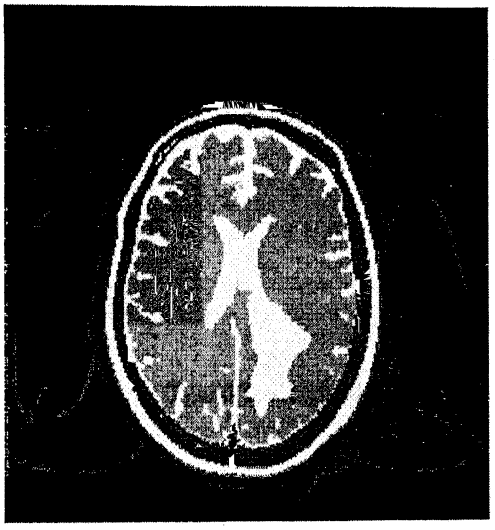
Figure 4.4 shows a segmented image obtained using Algorithm 4.3.1 on the test image set II, with B = 0.02, taking PD, T_2 and T_1 weighted images. It is observed that PD, T_2 and T_1 weighted images together produce good segmentation of different normal tissues along with the pathology present on the images. The segments

corresponding to CSF, white matter, gray matter and edema are well represented on this segmented image.



PD-T_2 weighted Images

PD-T_1 weighted Images



T_2-T_1 weighted Images

Figure 4.3: Segmented Images Obtained using Algorithm 4.3.1 to each pair of Test Image Set II

Also the segmentation is done using three weighted images for consecutive six cross-sections obtained from a single patient and the results are shown in Figure 4.5. All the results are obtained with B = 0.02, and the segmented images so obtained are

satisfactory. The number of iterations taken for cross-sections 1, 2, 3, 4, 5 and 6 are 47, 45, 34, 39, 50 and 59 respectively.

With the variation in the convergence criterion, i.e., varying the value of B from 0.05 to 0.30 with an increment of 0.05, the results so obtained are presented in Figure 4.6. The number of iterations taken in cases of B = 0.05, 0.10, 0.15, 0.20, 0.025 and 0.030 are 32, 14, 14, 14, 13 and 13, respectively. The sizes of major segments like white matter, edema, ventricles, etc. are almost similar in all these cases, but the graininess within the segments increases with the increase in the value of B.

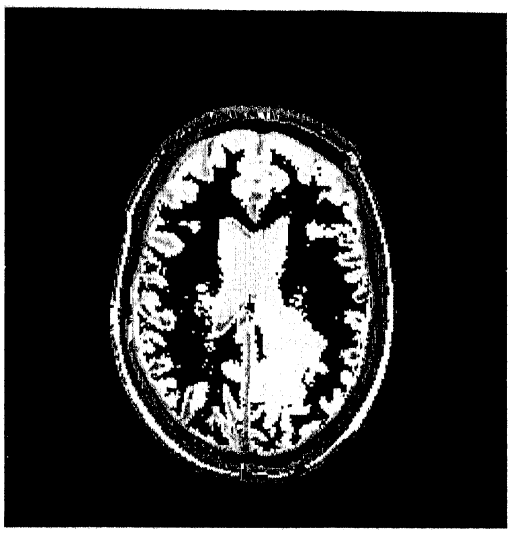
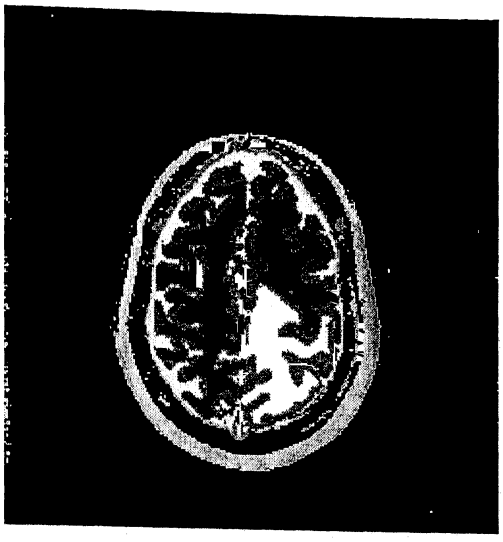
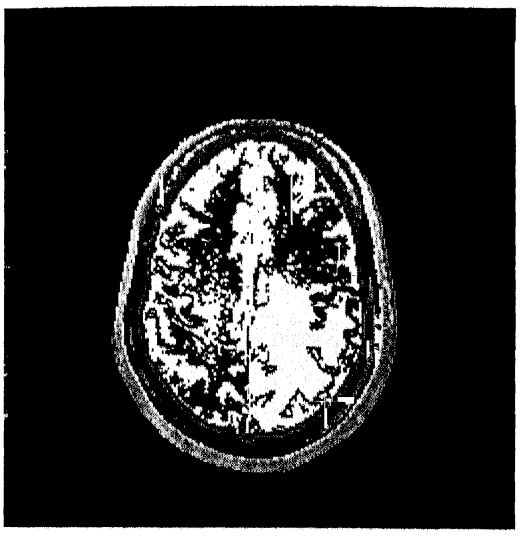


Figure 4.4: Segmented Image Obtained using Algorithm 4.3.1 for PD, T_2 and T_1 weighted Images of Test Image Set II

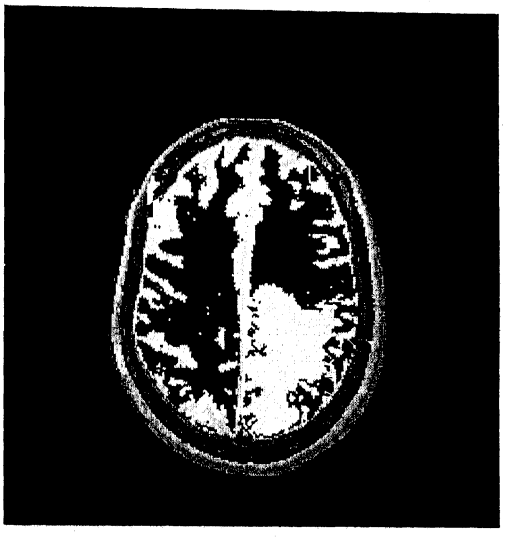
Cross-section 1



Cross-section 2

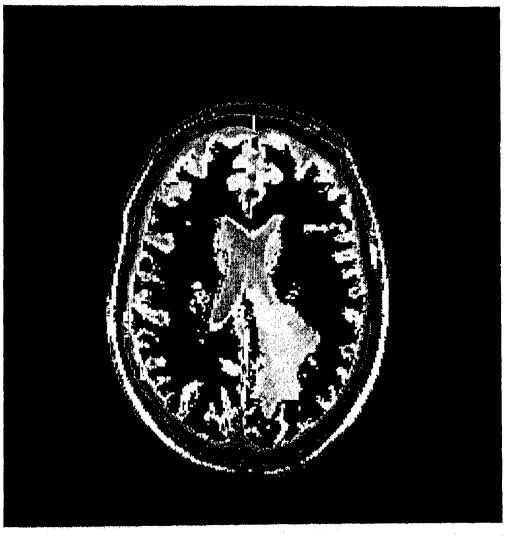


Cross-section 3



Cross-section 4

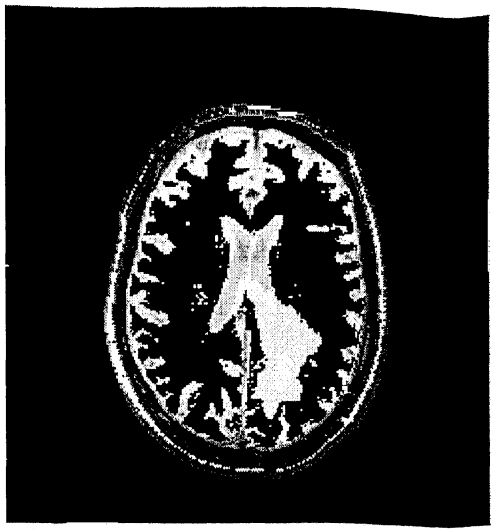
Cross-section 5



Cross-section 6

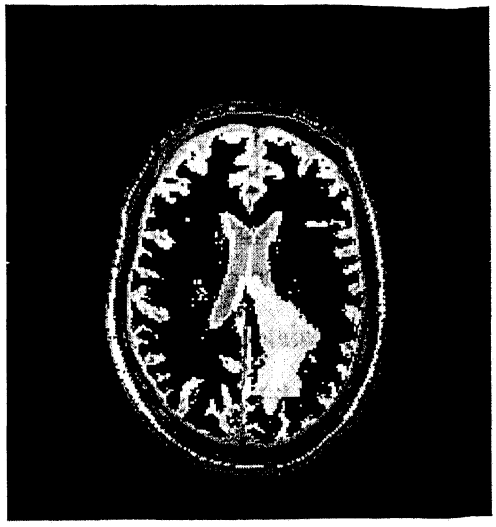
Figure 4.5: Segmented Images of Six Cross-sections in Case of Edema

B=0.05

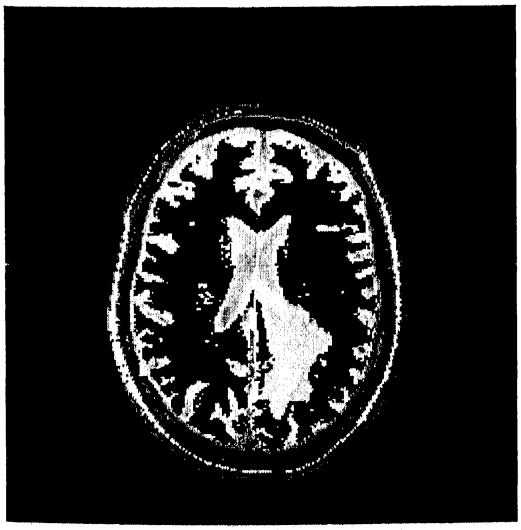


B=0.10

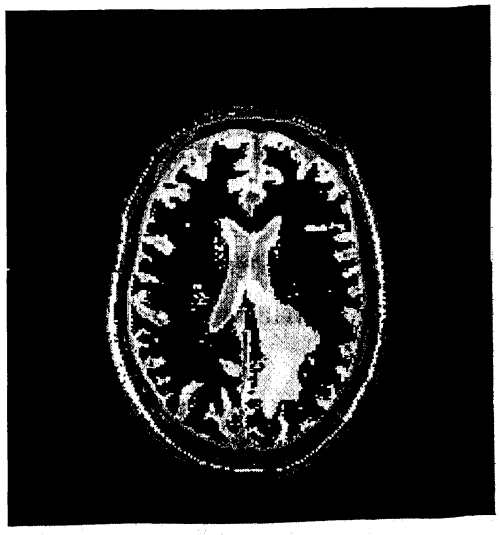
B=0.15



B=0.20



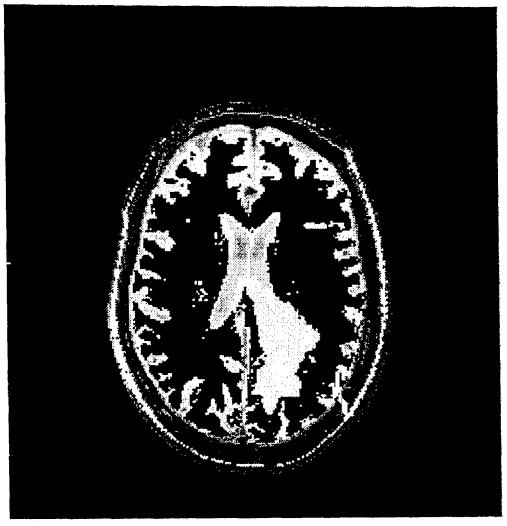
B=0.25



B=0.30

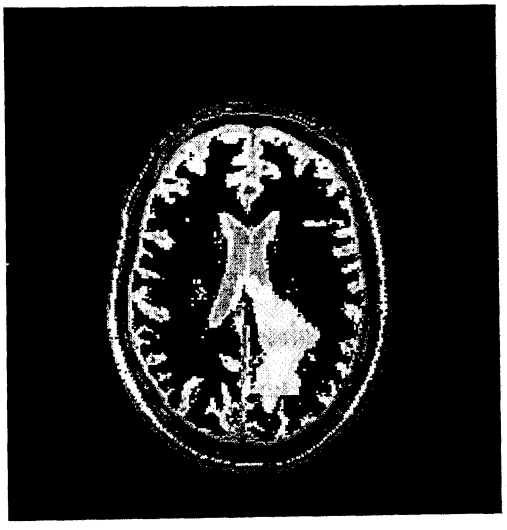
Figure 4.6: Segmented Images of Test Image Set II in Case of Edema with Different Values of B

B=0.05

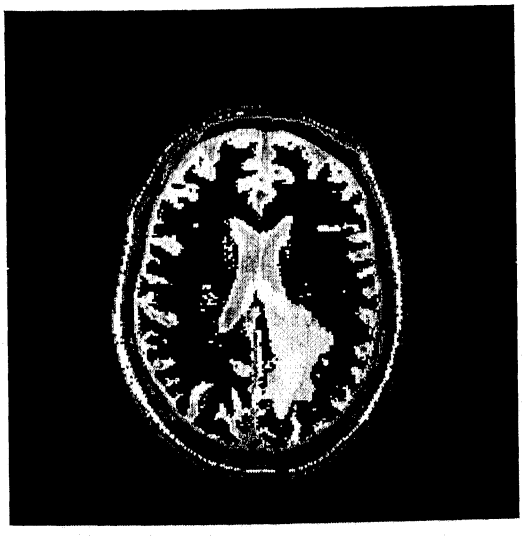


B=0.10

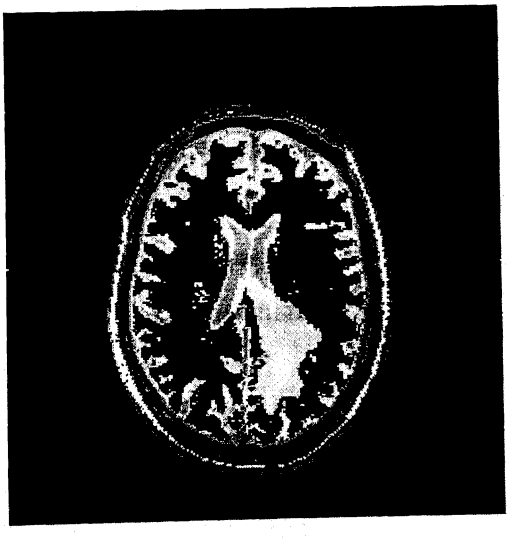
B=0.15



B=0.20



B=0.25



B=0.30

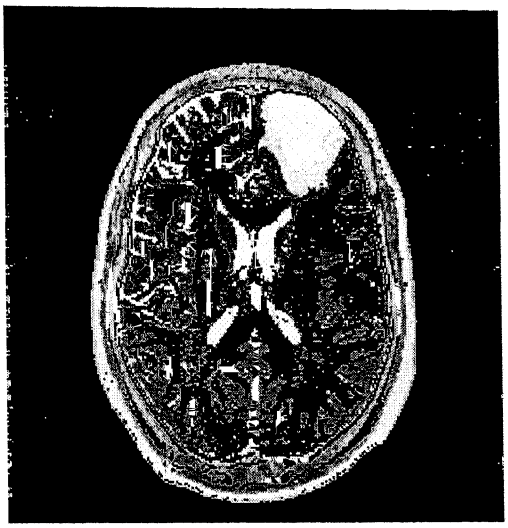
Figure 4.6: Segmented Images of Test Image Set II in Case of Edema with Different Values of B

4.5. OBSERVATIONS RELATED TO TEST IMAGE SET III

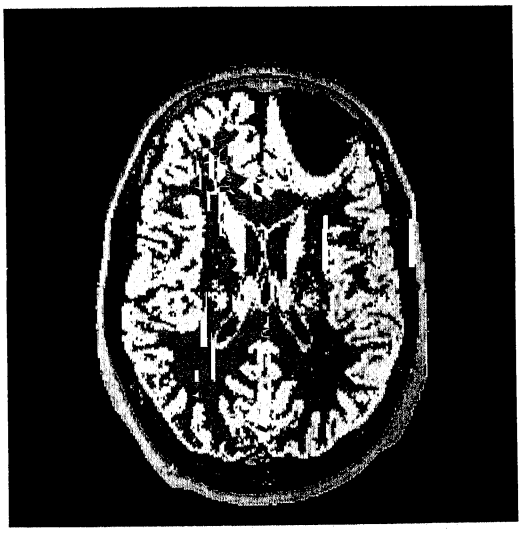
In test image set III, gliosis seems to be hyperintense on MT SE T_1-weighted image, and it is little visible on PD-weighted image. The contrast of either lesion or gliosis is not good on any of these weighted images. So, when we apply a segmentation algorithm based on MLE in case of a single image taking either PD, or MT SE T_1 weighted image, it gives single segment for gray matter and gliosis, where initial number of segments was fifteen. However, even with the increase in number of segments to more than twenty in the segmentation procedure, gliosis keeps on merging with the gray matter. This is because of the variation within the cranial part of the brain (consisting of skin, bone, muscle and fat) which is much more than that present between gray matter and gliosis, resulting in the further divisions of the segments within the cranial parts.

Using multi-resolution Algorithm 4.3.1 to all possible pairs of PD, T_2 and T_1 weighted images with B = 0.02, the results obtained are shown in Figure 4.7. In these cases, lesion and the perilesional gliosis are well separated in PD-T_2 pair, whereas white matter is segmented satisfactorily in PD-T_1 pair. The result obtained applying multi-resolution Algorithm 4.3.1 with B = 0.02, taking PD, T_2 and T_1 weighted images of test image set III, is also shown in Figure 4.7. The segmented image clearly shows the separate segments corresponding to lesion, perilesional gliosis and white matter. This has happened since the contrast between the lesion and the perilesional gliosis compared with the gray matter is high on T_2 weighted image due to T_2 abnormality of pathology. So, when we combine PD, T_1, and T_2 weighted images for segmentation, it is observed that all major brain components like gray matter, white matter, CSF and abnormalities (lesion and gliosis) have distinct normal distributions. The classifications of different tissues were based on Bayes' decision theory (R.O. Duda and P.E. Hart [1973]), according to which the pixel belongs to that segment for which its probability is maximum among all the segments.

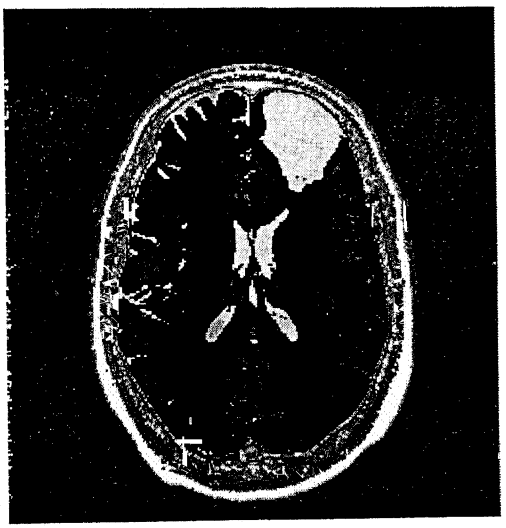
The regions of lesion and ventricles are represented by same gray level on segmented image as both of them together form a single distribution. So, the distinction of the lesion from ventricles is difficult on the basis of gray level and hence



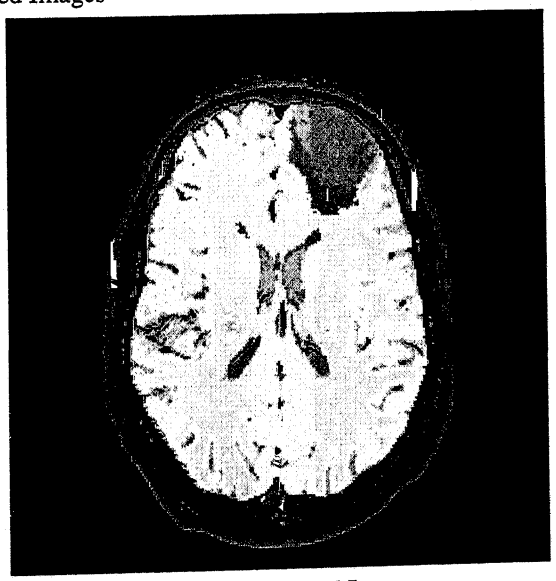
PD-T_2 weighted Images



PD-T_1 weighted Images



T_2-T_1 weighted Images

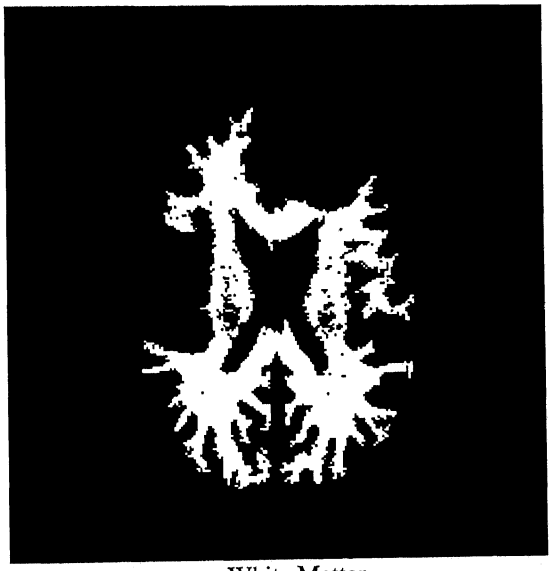


PD-T_2-T_1 weighted Images

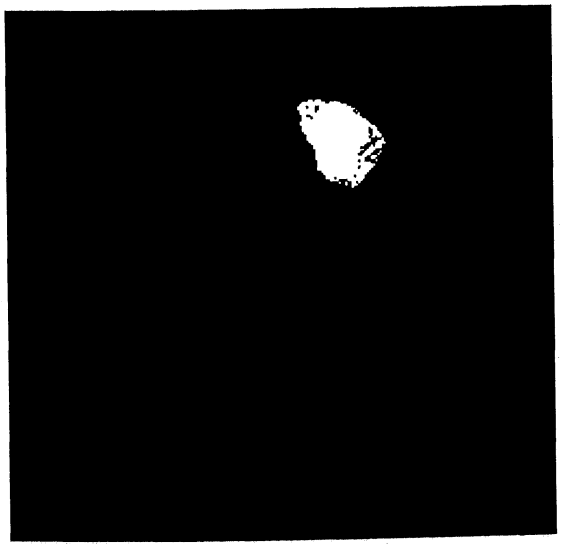
Figure 4.7: Segmented Images Obtained using Algorithm 4.3.1 to Test Image Set III

extraction of lesion has been done using seed - growing algorithm, with four or eight connectivity of pixels, having seed pixel from the region of the lesion. Similarly, also the white matter and the gliosis have been extracted out using seed growing algorithm with eight connectivity neighborhood. The extracted parts of the white matter, the lesion and the perilesional gliosis are shown in Figure 4.8.

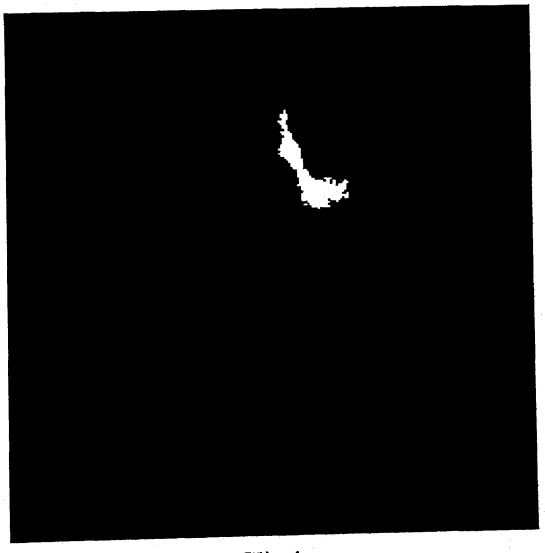
A class of different tissues may give one segment as in the case of lesion and cerebrospinal fluid present in ventricles, similarly one tissue may not make a single connected region like gray matter which comprises of several segments on an image. Two



White Matter



Lesion

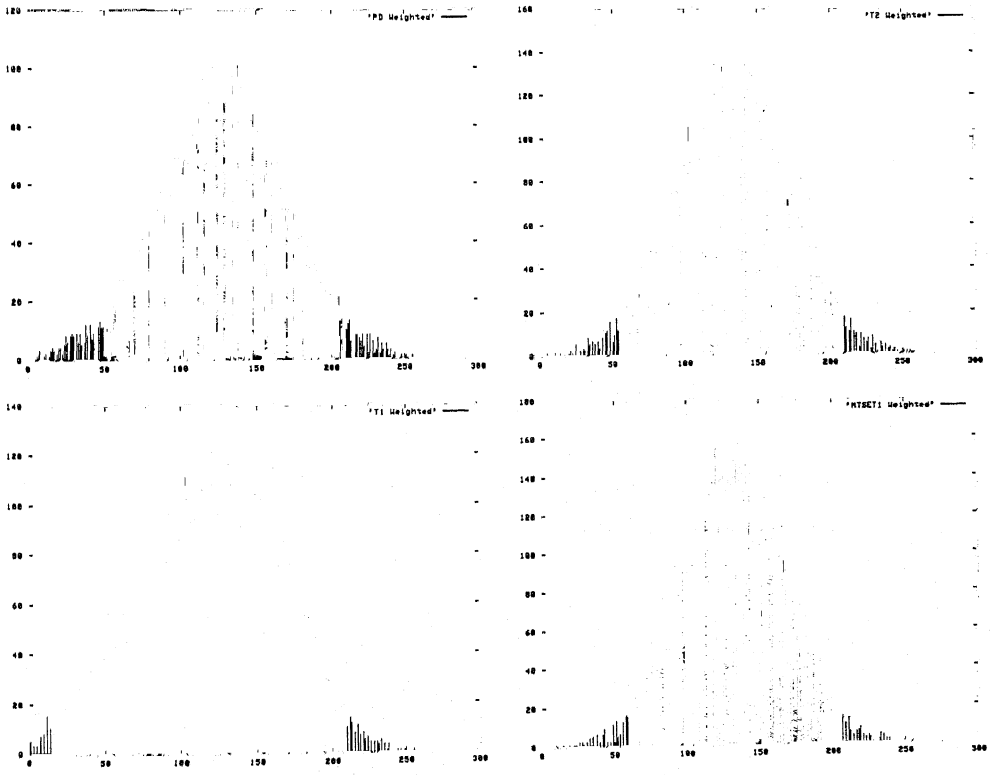


Gliosis

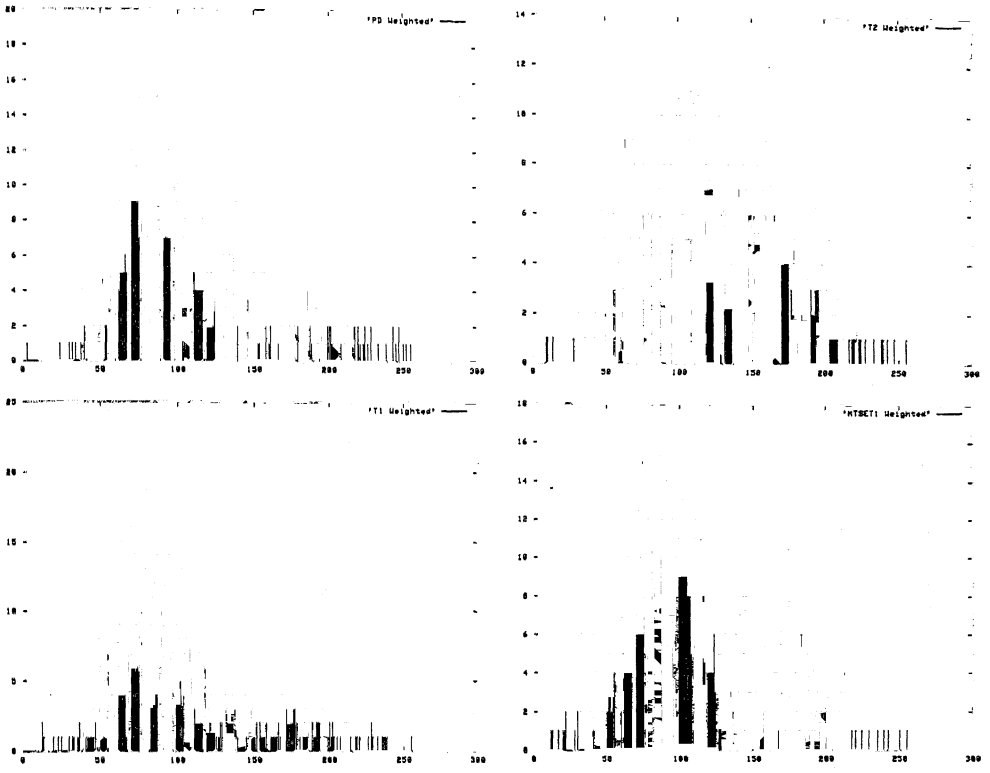
Figure 4.8: Extracted White Matter, Lesion and Perilesional Gliosis

or more tissues falling under a single distribution can be extracted and the structure of gray matter comprising of different components can be obtained on the basis of anatomical knowledge of brain and the connectivity of the region.

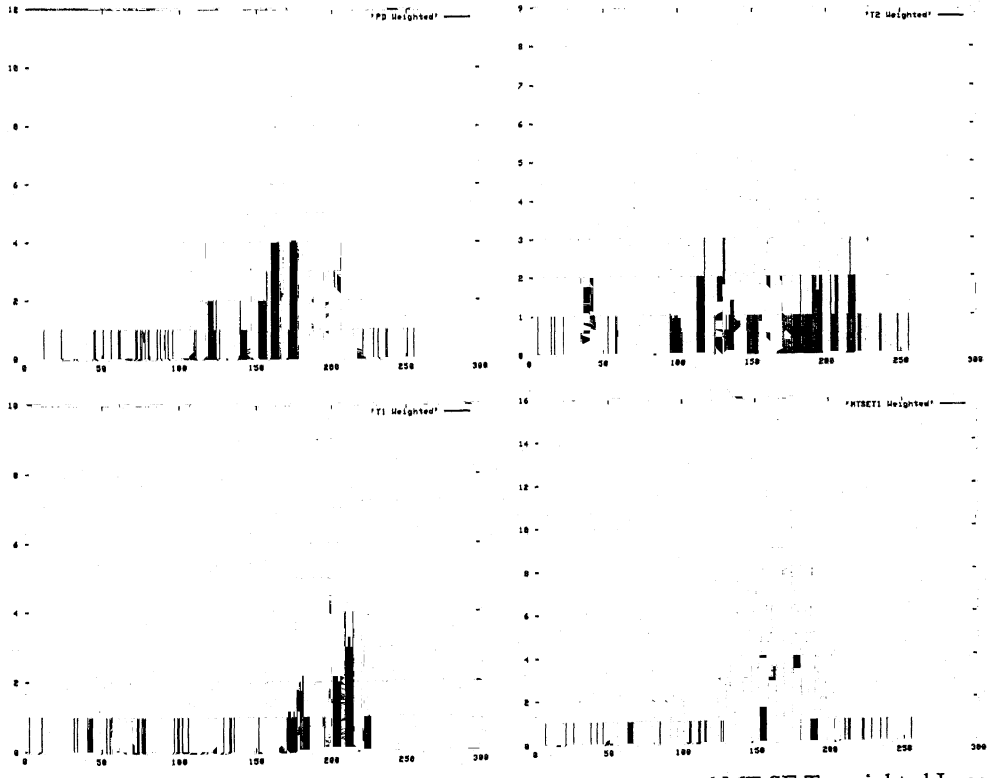
Graphs 4.1, 4.2 and 4.3 show the histogram plots of white matter, lesion and perilesional gliosis on PD, T_2, T_1 and MT SE T_1 weighted images. Distribution of white matter shows normal on all the weighted images, while that of lesion is normal on T_2 weighted image and close to normal on other weighted images. Distribution of perilesional gliosis is close to normal on PD and MT SE T_1 weighted images and two normal distributions are visible within gliosis on T_2 and T_1 weighted images. Gliosis shows two peaks of normal distributions on T_2 weighted image, as it is represented as a T_2 abnormality on T_2 weighted image. Similarly gliosis is not separately visible on T_1 weighted image as a separate entity.



Graph 4.1: Histogram Plots of White Matter in PD, T_2, T_1 and MT SE T_1 weighted Images



Graph 4.2: Histogram Plots of Lesion in PD, T_2, T_1 and MT SE T_1 weighted Images



Graph 4.3: Histogram Plots of Perilesional Gliosis in PD, T_2, T_1 and MT SE T_1 weighted Images

Means, variances and the correlation coefficients between the pairs of PD, T_2 and T_1 weighted images in cases of lesion, gliosis, and their contra-lateral regions have been obtained after segmentation and the results are shown in Tables 4.1 and 4.2. Results corresponding to white matter are separately given in Table 4.3. All the quantities are calculated on values normalized to 0-255.

Image	Lesion	Contra-lateral
PD weighted image ($\mu \pm \sigma$)	131.56 ± 7.54	145.53 ± 6.77
T_2 weighted image ($\mu \pm \sigma$)	192.03 ± 13.28	127.63 ± 21.77
T_1 weighted image ($\mu \pm \sigma$)	68.53 ± 6.89	106.76 ± 12.57
MT SE T_1 weighted image ($\mu \pm \sigma$)	64.85 ± 5.35	75.10 ± 4.50
Correlation coefficients (PD, T_2)	0.068290	-0.033513
Correlation coefficients (PD, T_1)	0.859128	0.251560
Correlation coefficients (T_1, T_2)	0.148065	-0.841327
No. of pixels	991	991

Table 4.1: Means, Variances, Correlation Coefficients of Lesion and its Contra-lateral part

Image	Gliosis	Contra-lateral
PD weighted image ($\mu \pm \sigma$)	153.75 ± 5.031	141.41 ± 5.07
T_2 weighted image ($\mu \pm \sigma$)	187.39 ± 23.93	120.82 ± 14.32
T_1 weighted image ($\mu \pm \sigma$)	102.54 ± 10.91	107.06 ± 10.49
MT SE T_1 weighted image ($\mu \pm \sigma$)	81.30 ± 4.47	74.47 ± 2.95
Correlation coefficients (PD, T_2)	0.028229	0.315359
Correlation coefficients (PD, T_1)	0.413414	-0.073068
Correlation coefficients (T_1, T_2)	-0.790365	-0.776473
No. of pixels	438	438

Table 4.2: Means, Variances, Correlation Coefficients of Perilesional Gliosis and its Contra-lateral part

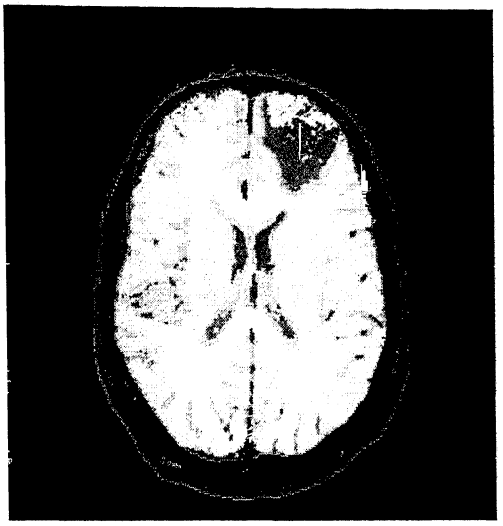
Image	White Matter
PD weighted image ($\mu \pm \sigma$)	136.80 ± 4.09
T_2 weighted image ($\mu \pm \sigma$)	111.94 ± 7.45
T_1 weighted image ($\mu \pm \sigma$)	112.01 ± 3.26
MT SE T_1 weighted image ($\mu \pm \sigma$)	73.45 ± 2.52
Correlation coefficients (PD, T_2)	0.259644
Correlation coefficients (PD, T_1)	0.397903
Correlation coefficients (T_1, T_2)	0.172379
No. of pixels	7557

Table 4.3: Means, Variances, Correlation Coefficients of White Matter

From the computed values for weighted images in contra-lateral regions corresponding to lesion and gliosis, it is clear that they correspond to similar tissues in that. The T_2 weighted values of lesion and gliosis are falling in one range, which is clear from the image also. Comparing the values in contra-lateral parts of lesion and gliosis and white matter, it becomes clear that the part of lesion along with gliosis falls within white matter.

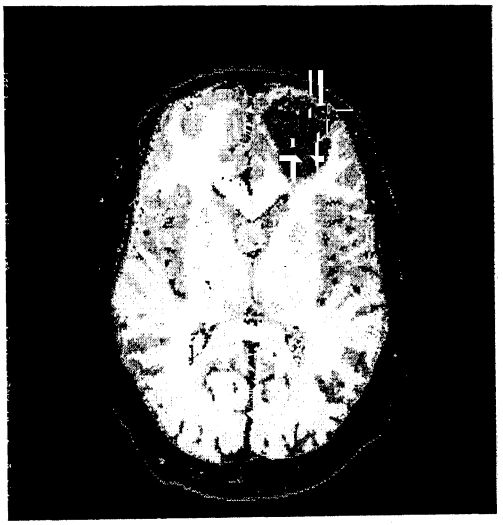
Segmentation using PD, T_2 and T_1 weighted images were obtained using Algorithm 4.3.1., with B = 0.02, for consecutive five cross-sections obtained from single patient, and the results are shown in Figure 4.9. The number of iterations taken for segmentation of cross-sections 1, 2, 3, 4 and 5 are 51, 59, 36, 45 and 43, respectively. The results corresponding to the variation in convergence criterion (varying the value of B from 0.05 to 0.30 with an increment of 0.05) are shown in Figure 4.10. The number of iterations taken in case of B = 0.05, 0.10, 0.015, 0.20, 0.25 and 0.30 are 38, 17, 16, 16, 15 and 12, respectively. The sizes of major segments like white matter, ventricles, etc., are almost similar in all these cases, but the graininess within the segments increases with the increase in the values of B. For values of B from 0.05 and above, the parts of both lesion and perilesional gliosis have been merged with gray matter.

Cross-section 1



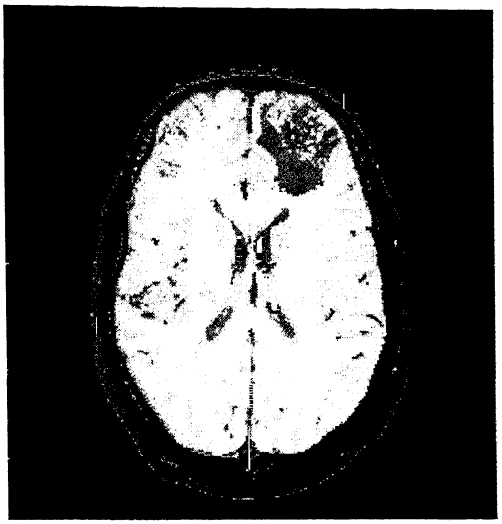
Cross-section 2

Cross-section 3



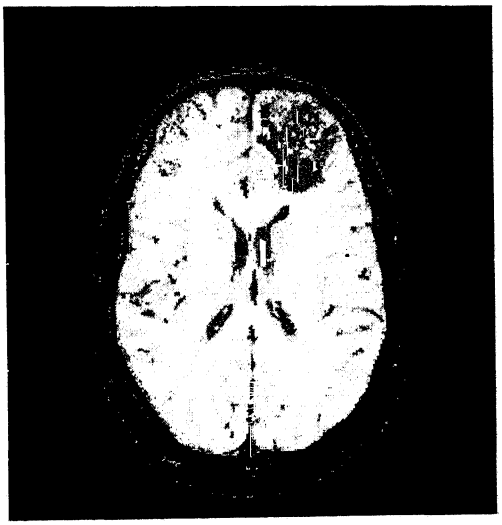
Cross-section 4

B=0.05



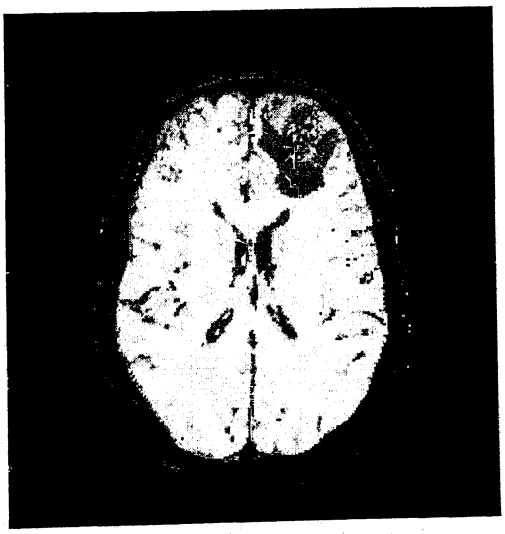
B=0.10

B=0.15



B=0.20

B=0.25



B=0.30

Figure 4.10: Segmented Images of Test Image Set III in case of Perilesional Gliosis with Different Values of B

In all our experiments, we have considered MR images of size 256x256. It is observed that the images reconstructed from 192x256 Fourier data are artificially extended to 256x256 grid by adding two strips of uniform distribution of zero values on both sides of the image. In such situations, Algorithms 4.2.1. and 4.3.1. were applied after the removal of these uniform distributions as both the algorithms would fail to give segments corresponding to uniform distributions due to zero variance in these distributions.

Segmentation results into two or more number of segments due to the presence of noise in the outside region of the brain (mostly air) on the images, while it seems to possess a single segment. This can be suppressed by thresholding the image and then removing it after getting a uniform distribution outside, but it alters other segments in the cranial part also, as the intensity of bone matches with that of exterior part on all the weighted images.

Sometimes segmentation leads to misclassification of certain pixels due to the inhomogeneity present within the tissues. This can be reduced by using a proper smoothing filter, which smoothens the region without disturbing the major edges before the application of segmentation algorithms or by considering further accuracy in the convergence criteria.

Our further experiments show that if one removes the cranial part of brain, then one would get a better segmentation as the inhomogeneity within each tissue in the interior part of the brain is less as compared to that of the cranial part. Also, one can study the cranial part of the brain separately on the basis of this algorithm.

4.6. BOUNDARY TRACING USING CUBIC SPLINES

In most of the magnetic resonance images, the boundary of any tissue or pathology seems to be blurred due to partial volume effect. The method proposed here is helpful in tracing the boundary of any closed region comprising of pathology or any normal tissue. The boundary of a region on an image is calculated from the minimum

energy convex combination of the original image, its histogram and the Sobel edge map of it. This combination sharpens edge so as to enable an easy determination of knot points on it for the construction of the cubic spline (S. Datta, R.K. Gupta, S.B. Rao, V.S.N. Kaliprasad and R.K.S. Rathore [2000]). The 2-D cubic spline interpolationg curve $z(t) = (x(t), y(t))'$ (J.H. Ahlberg, E.N. Nilson and J.L. Walsh [1967]) is obtained as

$$z(t) = \frac{M_{j-1}}{6h_j}\tau_j^3 + \frac{M_j}{6h_j}\tau_{j-1}^3 - \frac{1}{h_j}\left(z_{j-1} - \frac{M_{j-1}}{6}h_j^2\right)\tau_j + \frac{1}{h_j}\left(z_j - \frac{M_j}{6}h_j^2\right)\tau_{j-1}.$$

Here $\tau_j = t - t_j$, $h_j = t_j - t_{j-1}$, $j = 1,2,3, \ldots, N$. M_j and M_{j-1} equal $z''(t)$ at points $t = t_j, t_{j-1}$, where

$$\begin{pmatrix} 2 & \lambda_1 & 0 & \cdots & \cdots & 0 & 0 & \mu_1 \\ \mu_2 & 2 & \lambda_2 & \cdots & \cdots & 0 & 0 & 0 \\ 0 & \mu_3 & 2 & \cdots & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 2 & \lambda_{N-2} & 0 \\ 0 & 0 & 0 & \cdots & \cdots & \mu_{N-1} & 2 & \lambda_{N-1} \\ \lambda_N & 0 & 0 & \cdots & \cdots & 0 & \mu_N & 2 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ M_3 \\ \vdots \\ \vdots \\ M_{N-2} \\ M_{N-1} \\ M_N \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ \vdots \\ d_{N-2} \\ d_{N-1} \\ d_N \end{pmatrix},$$

$$d_j = \left(\frac{6}{h_j + h_{j+1}}\right)\left(\frac{z_{j+1} - z_j}{h_{j+1}} - \frac{z_j - z_{j-1}}{h_j}\right),$$

$$M_0 = M_N, \lambda_N = h_1/(h_N + h_1), \mu_N = 1 - \lambda_N,$$

$$\lambda_j = h_{j-1}/(h_j + h_{j-1}), \mu_j = 1 - \lambda_j, j = 1,2,\ldots,N-1.$$

A binary image is obtained after tracing the boundary using cubic spline, which is further segmented into different components on the basis of a seed growing algorithm. The area inside the spline can be calculated by conventional method by integrating cubic spline and also by counting the total number of pixels comprising of particular tissue.

The area inside the region is calculated by integrating the spline as:

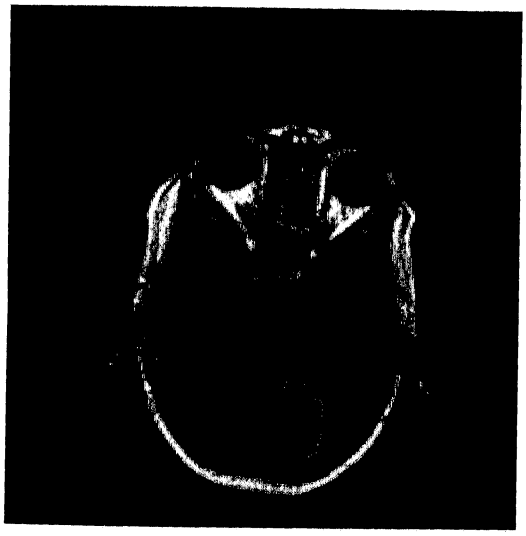
$$\sum_{j=1}^{N} \frac{h_j^2}{24} \nabla y_j (M_{x_j} + M_{x_{j-1}}) - \nabla x_j (M_{y_j} + M_{y_{j-1}}) + \frac{M_{y_{j-1}} M_{x_j} h_j^4}{720} - \frac{M_{x_{j-1}} M_{y_j} h_j^4}{720}$$

$$+ \sum_{j=1}^{N} \{(1/2)[(x_j - \xi)(y_{j-1} - \eta) - (x_{j-1} - \xi)(y_j - \eta)]\}$$

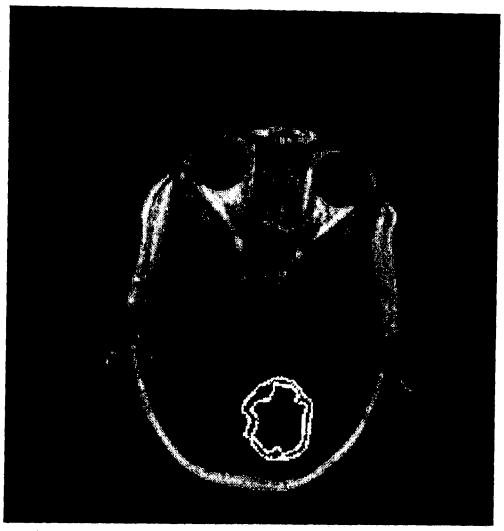
$$- \sum_{j=1}^{N} [\nabla x_j (y_{j-1} - \eta) - \nabla y_j (x_{j-1} - \xi)]$$

$$- \sum_{j=1}^{N} \{(1/2)[\nabla y_j (x_j + x_{j-1} - 2\xi) - \nabla x_j (y_j + y_{j-1} - 2\eta)]\},$$

where $\nabla y_j = y_j - y_{j-1}, \nabla x_j = x_j - x_{j-1}$. (ξ, η) is any point inside the region.



Original



Boundaries of Tuberculoma Superimposed on Original Image

Figure 4.11: Original Image and Boundaries of Tuberculoma Obtained using Cubic Splines Superimposed on Original Image

Area under	# Area	∫ Area
Outer curve	1101	1101.863
Inner curve	697	697.28418

Table 4.4: Total Area under Outer and Inner Rims of Tuberculoma

#Area by calculating the average of no. of pixels inside spline and no. of pixels inside spline along with the curve and ∫Area calculated by integrating spline curve.

This method is applied on an post contrast MT SE T_1 image which is obtained after injecting Gd-DTPA intravenously in the dose of 0.1 mmol/kg body weight. In this case the lesion is characterized by intracranial tuberculoma surrounded by edema. The inner and outer rims of tuberculoma are obtained using the proposed method. The areas calculated inside the rim by integrating the spline and also by discretely calculating the number of pixels within region are comparable. The result is summarized in Table 4.4 and the original image and the image having boundaries superimposed on it are shown in Figure 4.11. The inner and outer boundaries obtained here are very much close to the boundaries, which are seen on original image.

The method proposed here is useful in obtaining boundary of a lesion with discontinuous edges. This can be implemented for a boundary of any shape. Instead of storing all the boundary points, one can trace a complete boundary by storing knot points on the boundary.

4.7. CONCLUSIONS

It is not always that one has to use all of the three or more weighted images together to get a proper segmentation. This mainly depends on the region of interest and its contrast with respect to neighboring regions in each of these images, which suggests to select the number of images and the type of images to be used to get a proper segmentation. As we have discussed in section 4.4, to get the edema extracted from an image in case of image set I, it is sufficient to use only T_2 weighted image, while all of the PD, T_2 and T_1 weighted images of test image set II were used to segment the edema, irrespectively of that both the image sets are of same patient with different cross-sections. In second case, even two images are not good enough to segment the region of interest.

While in the case of post-traumatic epilepsy with perilesional gliosis, if one wishes to extract only gliosis part from an image, one can use only PD and T_2 weighted images and PD and T_1 weighted images are good enough for the segmentation of white matter. But, if we use PD, T_2 and T_1 weighted images for

segmentation, then extraction of gliosis, white matter, lesion, etc., are possible after application of multi-resolution segmentation algorithm.

The theoretical convergence of the MLE based iterative schemes is not known in the several cases that are possible. In all our experiments, nevertheless, we have observed the numerical convergence, except for the earlier mentioned cases of zero variances and with zero population size ratios, in 192×256 cases artificially extended to 256×256 grid.

REFERENCES

[01] Ahlberg, J.H., Nilson, E.N., and Walsh, J.L., "The theory of splines and their applications", Academic Press, 1967.

[02] Amartur, S.C., Piraino, D., and Takefuji, Y., "*Optimization neural networks for the segmentation of magnetic resonance images*", IEEE Trans. Medical Imaging, vol. MI-11, no. 2, pp. 215-220, 1992.

[03] Ashkar, G.P., and Modestino, J.W., "*The contour extraction problem with biomedical applications*", Computer Graphics and image Processing, vol. 7, pp. 331-355, 1978.

[04] Besag, J., "*Spatial interaction and the statistical analysis of lattice systems (with discussion)*", J. of Royal Statist. Soc., series B, vol. 36, no. 2, pp.192-326, 1974.

[05] Bojanic, R., and Shisha, O., "*Approximation of continuous, periodic functions by discrete linear positive operators*", Jr. Approx. Th., vol. 11, pp. 231-235, 1974.

[06] Bomans, M., Höhne, K-H., Tiede, U., and Riemer, M., "*3-D segmentation of MR images of the head for 3-D display*", IEEE Trans. Medical Imaging, vol. MI-9, pp. 177-183, 1990.

[07] Butzer, P.L., and Berens, H., "Semi-groups of Operators and Approximation", Springer Verlag, Berlin, 1967.

[08] Chakraborty, A., Staib, L.H., and Duncan J.S., "*Deformable boundary finding in medical images by integrating gradient and region information*", IEEE Trans. Medical Imaging, vol. MI-15, no. 6, pp. 859-870, 1996.

[09] Clark, M.C., Hall, L.O., Goldgof, D.B., Velthuizen, R., Murtagh F.R., and Silbiger, M.S., "*Automatic tumor segmentation using knowledge-based techniques*", IEEE Trans. Medical Imaging, vol. MI-17, no. 2, pp. 187-201, 1998.

[10] Clarke, L.P., Velthuizen, R.P., Camacho, M.A., Heine, J.J., Vaidyanathan, M., Hall, L.O., Thatcher, R.W., and Silbiger, M.L., "*MRI segmentation: Methods and applications*", Magnetic Resonance Imaging, vol. 13, no. 3, pp. 343-368, 1995.

[11] Clarke, L.P., Velthuizen, R.P., Phuphanich, S., Schellenberg, J.D., Arrington, J.A., and Silbiger, M., "*MRI: Stability of three supervised segmentation techniques*", Magnetic Resonance Imaging, vol. 11, pp. 95-106, 1993.

[12] Datta, S., Gupta, R.K., Rao, S.B., Kaliprasad, V.S.N., and Rathore, R.K.S., "*Segmentation of Lesions in MRI using Cubic Splines*", National Symposium on Magnetic Resonance and Biomolecular Structure and Function, TIFR India, January 17-20, 2000.

[13] DeVore, R.A., "The Approximation of Continuous Functions by Positive Linear Operators", Lecture Notes, Springer Verlag, Berlin-NewYork, 1972.

[14] Duda, R.O., and Hart, P.E., "Pattern Recognition and Scene Analysis", John Wiley, New York, 1973.

[15] Geman, S., and Geman, D., "*Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-6, no. 6, pp. 721-741, 1984.

[16] Gonzalez, R.C., and Woods, R.E., "Digital Image Processing", Addison Wesley, 1999.

[17] Haralick, R.M., "*Digital step edges from zero crossing of second directional derivatives*", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-6, no. 1, pp. 58-68, 1984.

[18] Haralick, R.M., Sternberg, S.R., and Zhuang, X., "*Image analysis using mathematical morphology*", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-9, no. 4, pp. 532-550, 1987.

[19] Hildreth, E.C., "*The detection of intensity changes by computer and biological vision systems*", Comput. Vision, Graphics, Image Processing, vol. 22, pp. 1-27, 1983.

[20] Horowitz, A.L., "MRI Physics for Physicians", Springer Verlag, New York, 1989.

[21] Huertas, A., and Medioni, G., "*Detection of intensity changes with subpixel accuracy using Laplacian-Gaussian masks*", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-8, no. 5, pp. 651-664, 1986.

[22] Jain, A.K., "Fundamentals of Digital image Processing", Prentice-Hall of India, New Delhi, 1995.

[23] Kaliprasad, V.S.N., "Iterative De-blurring in Magnetic Resonance Imaging", Ph.D. Thesis, I.I.T. Kanpur, 2000.

[24] Korovkin, P.P., "Linear Operators and Approximation Theory", Hindustan Publ. Corp., Delhi, 1960.

[25] Lundervold, A., and Storvik, G., "*Segmentation of brain parenchyma and cerebrospinal fluid in multispectral magnetic resonance images*", IEEE Trans. Medical Imaging, vol. MI-14, no. 2, pp. 339-349, 1995.

[26] Luo, D., "Pattern Recognition and Image Processing", Horwood Publishing, Chichester, 1998.

[27] Marr, D., and Hildreth, E.C., "*Theory of edge detection*", Proc. Roy. Soc. London, B, vol. 207, pp. 187-217, 1980.

[28] Montanari, U., "*On the optimal detection of curves in noisy pictures*", Commun. ACM, vol. 14, pp. 335-345, 1971.

[29] Natanson, I.P., "Constructive Function Theory, volume I", Frederick Ungar Publ. Co., New York, 1964.

[30] Partain, C.L., Price, R.R., Patton, J.A., Kulkarni, M.V., and James Jr., A.E., "Magnetic Resonance Imaging" volume II, W.B. Saunders Company, Philadelphia, 1988.

[31] Pham, D.L., and Prince, J.L., "*Adaptive fuzzy segmentation of magnetic resonance images*", IEEE Trans. Medical Imaging, vol. 18, no. MI-9, pp. 737-752, 1999.

[32] Rajapakse, J.C., Giedd, J.N., and Rapoport, J.L., "*Statistical approach to segemntation of single-channel cerebral MR images*", IEEE Trans. Medical Imaging, vol. MI-16, no. 2, pp. 176-186, 1997.

[33] Rao, C.R., "Linear Statistical Inference and Its Applications", Wiley Eastern Limited, 1989.

[34] Rathore, R.K.S., "*A generalized frame-work for inverse and equivalence results in approximation theory*", pp. 20-36 (in Mathematical Analysis and Applications, ed. A.P. Dwivedi), Narosa Publishing House, New Delhi, 2000.

[35] Rathore, R.K.S., "Approximation of Unbounded Functions by Linear Positive Operators", D.Sc. Thesis, Delft University, 1974a.

[36] Rathore, R.K.S., "Linear Combination of Linear Positive Operators and Generating Relations in Special Functions", Ph.D. Thesis, I.I.T. Delhi, 1973.

[37] Rathore, R.K.S., "*On a sequence of linear trigonometric polynomial operators*", SIAM J. Math. Anal., vol. 5, pp. 908-917, 1974b.

[38] Rathore, R.K.S., Datta, S., Gupta, R.K., Rao, S.B., and Verma, R., "*An MLE based segmentation method for quantitation in MR images*", 9th Scientific Meeting of ISMRM, Glasgow, April 21-27, 2001.

[39] Reddick, W.E., Glass, J.O., Cook, E.N., Elkin, T.D., and Deaton, R.J., "*Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks*", IEEE Trans. Medical Imaging, vol. MI-16, no. 6, pp. 911-918, 1997.

[40] Russ, J.C., "The Image Processing Handbook", CRC Press, 1992.

[41] Schoenberg, I.J., "*On variation diminishing approximation methods*", pp. 249-274, (in On Numerical Approximation, ed. R.E. Langer), The University of Wisconsin Press, Madison, 1959.

[42] Serra, J., "Image Analysis and Mathematical Morphology", Academic Press, London, 1982.

[43] Shapiro, H.S., "Smoothing and Approximation of Functions", Van Nostrand, New York, 1969.

[44] Sikkema, P.C., and Rathore, R.K.S., "*Convolutions with powers of bell-shaped functions (report)*", Onderafdeling der Wiskunde Technische Hogeschool Delft, 1976.

[45] Staib, L.H., and Duncan, J.S., "*Boundary finding with parametrically deformable models*", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-14, no. 11, pp. 1061-1075, 1992.

[46] Torre, V., and Poggio, T., "*On edge detection*", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-8, no. 2, pp. 147-163, 1986.

[47] Vaidyanathan, M., Clarke, L.P., Heidtman, C., Velthuizen, R.P., and Hall, L.O., "*Normal brain volume measurements using multispectral MRI segmentation*", Magnetic Resonance Imaging, vol. 15, no. 1, pp. 87-97, 1997.

[48] Wells III, W.M., Grimson, W.E.L., Kikinis, R., and Jolesz, F.A., "*Adaptive segmentation of MRI data*", IEEE Trans. Medical Imaging, vol. MI-15, no. 4, pp. 429-442, 1996.

[49] Yoshida, K., "Functional Analysis", Springer-Verlag, New-York, 1971.

[50] Zhang, Y., Brady, M., and Smith, S., "*Segmentation of brain MR images through a hidden markov random field model and the expectation maximization algorithm*", IEEE Trans. Medical Imaging, vol. MI-20, no. 1, pp. 45-57, 2001.